blackwell great minds

kant

allen w. wood

Copyrighted material

kant

blackwell great minds
edited by Steven Nadler

The *Blackwell Great Minds* series gives readers a strong sense of the fundamental views of the great western philosophers and captures the relevance of these philosophers to the way we think and live today.

1. Kant by Allen W. Wood

Forthcoming
Aristotle by Jennifer Whiting
Augustine by Gareth Matthews
Descartes by Andre Gombay
Nietzsche by Richard Schacht
Plato by Paul Woodruff
Sartre by Katherine J. Morris
Spinoza by Don Garrett
Wittgenstein by Hans Sluga

blackwell great minds

kant

allen w. wood

© 2005 by Allen W. Wood

BLACKWELL PUBLISHING
350 Main Street, Malden, MA 02148-5020, USA
108 Cowley Road, Oxford OX4 1JF, UK
550 Swanston Street, Carlton, Victoria 3053, Australia

The right of Allen W. Wood to be identified as the Author of this Work has been asserted in accordance with the UK Copyright, Designs, and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs, and Patents Act 1988, without the prior permission of the publisher.

First published 2005 by Blackwell Publishing Ltd

Library of Congress Cataloging-in-Publication Data

Wood, Allen W.
Kant / Allen W. Wood.
p. cm. — (Blackwell great minds ; 1)
Includes bibliographical references and index.
ISBN 0-631-23281-8 — ISBN 0-631-23282-6 (pbk.)
1. Kant, Immanuel, 1724–1804. I. Title. II. Series.

B2798.W88 2005
193—dc22
2003026374

A catalogue record for this title is available from the British Library.

Set in 9.5/12pt Trump Mediaeval
by Graphicraft Limited, Hong Kong
Printed and bound in the United Kingdom
by MPG Books Ltd, Bodmin, Cornwall

The publisher's policy is to use permanent paper from mills that operate a sustainable forestry policy, and which has been manufactured from pulp processed using acid-free and elementary chlorine-free practices. Furthermore, the publisher ensures that the text paper and cover board used have met acceptable environmental accreditation standards.

For further information on
Blackwell Publishing, visit our website:
http://www.blackwellpublishing.com

To

Robert Merrihew Adams

contents

preface viii
abbreviations xiii

1 life and works 1
2 synthetic *a priori* cognition 24
3 the principles of possible experience 46
4 the limits of cognition and the ideas of reason 63
5 the transcendental dialectic 84
6 philosophy of history 110
7 ethical theory 129
8 the theory of taste 151
9 politics and religion 171

index 189

preface

The aim of this book is to introduce the philosophical thought of Immanuel Kant, especially to readers who are not yet familiar with it. Scholarly discussions and footnotes, therefore, have been kept to an absolute minimum. I have included some references to Kant's writings, but no more than I thought was minimally necessary to document my claims about what Kant says and enable the reader to look at the evidence in its proper context. The literature on Kant is vast, and much of it is of very high philosophical as well as scholarly quality. At the end of each chapter there are recommendations for further reading, aimed broadly at recommending the best books on the topics discussed in that chapter. They are in no sense bibliographies claiming to be complete or even particularly representative of the literature. The books I have recommended are among those I think are best, but the recommendations are also biased toward recent literature, since bibliographies of older literature are readily available (for example, in Paul Guyer (ed.) *The Cambridge Companion to Kant*; New York, 1992).

What is most remarkable about the philosophy of Kant, in my opinion, is the wide range of topics on which his thoughts repay careful study. In so many areas – not only in metaphysics but in natural science, history, morality, the critique of taste – he seems to have gone to the root of the matter, and at least raised for us the fundamental issues, whether or not we decide in the end that what he said about them is correct. In his brief, five-page essay on the question "What is Enlightenment?" for example, he locates the essence of enlightenment not in learning or the cultivation of our intellectual powers but in the courage and resolve to think for oneself, to emancipate oneself from tradition, prejudice, and every form of authority that offers us the comfort and security of letting someone else do our thinking for us. Kant's essay enables us to see that the issues raised by the challenge of the Enlightenment are still just as much with us as they were in the eighteenth century.

In a short book that attempts to cover the entire thought of such a wide-ranging philosopher, some things of importance are unavoidably omitted or slighted. Some things – notably, Kant's philosophy of natural

science and his ethical outlook – are much more important to his philosophy than the space devoted to them in this book would suggest. About half of the following book (chapter 2 though chapter 5) deals with the *Critique of Pure Reason* – Kant's longest published work, also his most famous and arguably his greatest lasting contribution to philosophy. But I have also devoted proportionally more space to Kant's theoretical philosophy than I might have because I have already written about Kant's practical (or moral) philosophy extensively elsewhere, especially in *Kant's Ethical Thought* (New York, 1999). Some of the basic ideas in Kant's theory of the physical sciences are discussed in chapters 2 and 3, but a proper appreciation of Kant's philosophy would require a more extensive treatment than I can provide here of the *Metaphysical Foundations of Natural Science* (1786). Some new ideas about the relation of philosophy to physics, strikingly different from anything he had written previously, were also the focus of Kant's final, tentative thoughts in a fragmentary and incomplete work known (since early in the twentieth century) as the *Opus Postumum*. In this book I have hardly mentioned that final bold adventure in thinking on which Kant had embarked just before his mental powers were cruelly ravaged by old age and then silenced forever by death. For those who do want to explore this final phase of Kant's philosophical thinking, I recommend the books by Eckart Förster and Michael Friedman listed under "Further Reading" at the end of chapters 1 and 2 respectively.

The first chapter of this book is biographical. This way of beginning a book about a philosopher is, however, highly questionable, and calls for some discussion. I begin with Kant's life because someone studying Kant for the first time, especially someone who is a relative beginner to the study of philosophy itself, will probably have an initial curiosity about who he was and how he lived. This is perfectly understandable and healthy. Yet those who have studied philosophy and its history soon learn that familiarity with the character or personality of philosophers is seldom very helpful in understanding their contributions to philosophy. Kant's life is of authentic interest to those of us who study his philosophy because it helps us to understand his world, both intellectual and material, and the relatively immediate aims, personal or social, which may have influenced his thoughts. Knowing about this may help us to understand why he thought and said some of the things he did, and therefore aid us in interpreting his ideas. Beyond that, our interest in his life may be historical, or antiquarian, or it may be mere idle curiosity. But it has nothing at all to do with his philosophy.

Especially to be avoided is approaching Kant's life in a spirit of hero worship or hagiography – as though our interest in a philosopher's thoughts is, or ought to be, proportional to our admiration for the thinker as a human being. If there have been any true saints or heroes

among important figures in the history of philosophy, we would do well entirely to ignore their heroism and saintliness in studying their philosophical thoughts. It is unhealthy and completely unphilosophical to venerate philosophers of the past as gurus at whose feet we should sit in order to absorb their wisdom. Such an attitude toward any other person, whether living or dead, betrays a contemptible slavishness of mind that it is incompatible with doing philosophy at all. In holding this opinion, I am, incidentally, also being a good Kantian, since Kant regarded the practice of those who set up others as models for imitation as morally corrupt, tending sooner to produce either self-contempt or envy than virtue. But that is all the more reason to apply Kant's view on this matter to Kant himself. Even the view itself should be given no credit at all just because Kant held it, but should be held only because experience shows it to be true – and true even about Kant himself.

It is a sometimes uncomfortable fact that the philosophers of the past whose thoughts we study with most profit were not especially fine human beings. The only way to deal with this fact is to face up squarely to the cognitive dissonance it occasions and then to resolve to set it aside as irrelevant to anything that could be of legitimate interest in deciding which philosophers to study. If a past philosopher, Kant for instance, was an admirable person, that still gives us no reason to study his philosophical thoughts if they were unoriginal or mediocre and do not repay our careful investigation and critical reflection. If the philosopher was a thoroughly unattractive character, or even if some of his opinions on morality or politics offend enlightened people today, it may still be true that his contributions to philosophy are indispensable to our understanding of philosophical problems and of the history of people's reflections on them. If we study the writings of the admirable philosopher in order to honor his virtuous character, then we are merely wasting time and effort that could have been better employed. By the same token, if we refuse to study the writings of the personally repulsive philosopher either because we think our neglect justly punishes him for his misdeeds or his evil opinions, or because we want to avoid being influenced by such a pernicious character, then all we accomplish by this foolish exercise in self-righteousness and closed-mindedness is to deprive ourselves of what we might have learned both from attaining to his insights and from exposing his errors. It is always sad to see philosophy students, and sometimes even professional philosophers, missing out on many things they might have learned on account of their moral or political approval or disapproval of the personality or opinions of some long-dead philosopher, who is far beyond their poor power to reward or punish. The only people we punish in this way are ourselves, and also those around us, or in the future, whom we might have influenced for the better if we had educated ourselves more wisely.

In Kant's case, I do not think that he was either a particularly admirable or a particularly unlikable human being. Rather, like most human beings, especially interesting ones, his character contained a rich mixture of attractive and unattractive traits. He was hard-working, patient, and utterly devoted to his work as a scientist, scholar, and philosopher, but he was also both shrewd and ambitious, never missing out on the personal advantages he gained through the professional success and prosperity he eventually achieved. He was a gregarious, sociable man, but sometimes quarreled with his friends and a number of his friendships came to an abrupt end. Though Kant believed above all in thinking for oneself, in his habits and lifestyle he seems at times to have been curiously open to the influence of certain friends – early in life, to Johann Daniel Funk, later in life to Joseph Green. He had a fierce love of the search for truth and of independent thinking, but he could also be jealous of his reputation, and mean-spirited toward students or followers he thought had personally betrayed him. He was not always above the intellectual cliquishness and academic backbiting characteristic of his time (and of intellectuals and academics of any time). Kant was a partisan of liberal reforms in education and especially in religion. He was a proponent of republicanism in politics, and of the proposition that states should relinquish some of their sovereign independence to a world-federation in the interests of international peace and the progressive development of the human species. He also uncompromisingly condemned European imperialism in other parts of the world, regarding all the pretended attempts of Europeans to 'liberate' or 'civilize' others as inherently unjust and hypocritical. But he also fully accepted and advocated the inferior status of women in society, and he held some views about non-European cultures and peoples that can be described only as racist. On the whole, Kant's was among the most progressive minds of his age in social and political matters. Yet some of his opinions on moral and political issues are either shocking or laughable to all enlightened people today. Rather than taking that as an occasion for venomous thoughts against Kant, we would be wiser to see it as a measure of the success of minds like his, philosophers who hoped they could promote better ways of thinking for the future, even if that might include the rejection of some of their own dearly held opinions. Whatever Kant's errors or vices, we would most definitely not be wrong in thinking of him as a philosopher for whom such hopes were an important spring of his own philosophical activity.

It is of course relevant to evaluating Kant's philosophy what his opinions were. But we are guaranteed to learn nothing from studying philosophy if we approach the writings of philosophers with the sole aim of trying to decide the extent to which the views expressed in them are in agreement with what we have decided beforehand that all people of good

will must believe. If this is the only spirit in which you can read works in the history of philosophy, then both you and the world at large would be better off if you simply remained ignorant of the history of philosophy and did not put on a show of knowing anything about it.

The true measure of Kant's value as an object of study by philosophers is the richness of the thoughts we have when we make the attempt to understand and also critically evaluate what he wrote and thought, and to relate those thoughts and our critical reflections on them to the philosophical problems that still occupy us. By that measure, to those who know him Kant is among the greatest philosophers who ever lived, whatever sort of man he may have been, and whatever we may think of his opinions on topics we care about.

I will also admit that the boldness of Kant's insights and the power of his arguments sometimes awaken in me feelings of admiration toward him. If I have been successful in presenting Kant in this book, then my exposition may perhaps awaken such feelings toward him in my readers as well. Anticipating the possibility of such success, I therefore issue the following advice, drawn from my own experience: When I find myself beginning to read Kant, or any philosopher, in a spirit of *veneration*, then that's a sign that I should stop reading him for a while and choose instead the writings of some other great philosopher (Hume, say, or Hegel) regarding whom such exceedingly anti-philosophical sentiments are not presently sapping my critical powers and clouding my good judgment.

A.W.W.
Stanford, California
July, 2003

abbreviations

Ak	*Immanuel Kants Schriften.* Ausgabe der königlich preussischen Akademie der Wissenschaften (Berlin: W. de Gruyter, 1902–). Unless otherwise noted, writings of Immanuel Kant will be cited by volume:page number in this edition.
Ca	*Cambridge Edition of the Writings of Immanuel Kant* (New York: Cambridge University Press, 1992–). This edition provides marginal Ak volume:page citations. Specific works will be cited using the following system of abbreviations (works not abbreviated below will be cited simply as Ak volume:page).
EF	*Zum ewigen Frieden: Ein philosophischer Entwurf* (1795), Ak 8 *Toward perpetual peace: A philosophical project*, Ca Practical Philosophy
G	*Grundlegung zur Metaphysik der Sitten* (1785), Ak 4 *Groundwork of the metaphysics of morals*, Ca Practical Philosophy
I	*Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht* (1784), Ak 8 *Idea for a universal history with a cosmopolitan aim*, Ca Anthropology, History and Education
KrV	*Kritik der reinen Vernunft* (1781, 1787). Cited by A/B pagination (according to the convention, adopted in the twentieth century, of referring to the first edition as 'A' and the second edition as 'B'). *Critique of pure reason*, Ca Critique of Pure Reason
KpV	*Kritik der praktischen Vernunft* (1788), Ak 5 *Critique of practical reason*, Ca Practical Philosophy
KU	*Kritik der Urteilskraft* (1790), Ak 5 *Critique of the power of judgment*, Ca Critique of the Power of Judgment
MA	*Mutmaßlicher Anfang der Menschengeschichte* (1786), Ak 8 *Conjectural beginning of human history*, Ca Anthropology, History and Education

MS	*Metaphysik der Sitten* (1797–8), Ak 6 *Metaphysics of morals*, Ca Practical Philosophy
O	*Was heißt: Sich im Denken orientieren?* (1786), Ak 8 *What does it mean to orient oneself in thinking?* Ca Religion and Rational Theology
P	*Prolegomena zu einer jeden künftigen Metaphysik* (1783), Ak 4 *Prolegomena to any future metaphysics*, Ca Theoretical Philosophy after 1781
R	*Religion innerhalb der Grenzen der bloßen Vernunft* (1793–4), Ak 6 *Religion within the boundaries of mere reason*, Ca Religion and Rational Theology
SF	*Streit der Fakultäten* (1798), Ak 7 *Conflict of the faculties*, Ca Religion and Rational Theology
TP	*Über den Gemeinspruch: Das mag in der Theorie richtig sein, taugt aber nicht für die Praxis* (1793), Ak 8 *On the common saying: That may be correct in theory but it is of no use in practice*, Ca Practical Philosophy
VA	*Anthropologie in pragmatischer Hinsicht* (1798), Ak 7 *Anthropology from a pragmatic standpoint*, Ca Anthropology, History and Education
WA	*Beantwortung der Frage: Was ist Aufklärung?* (1784), Ak 8 *An answer to the question: What is enlightenment?* Ca Practical Philosophy

chapter 1

life and works

The philosophical thought of Immanuel Kant marks the division between the two main periods in the history of modern philosophy. Retrospectively, Kant's philosophy was the last great attempt to solve the problems that had occupied philosophers in the seventeenth and eighteenth centuries. These included providing a philosophical foundation for the new science, working out the relation of this new view of nature to the traditional conceptions of metaphysics, morality, and religion, and defining the limits of our capacities to know both natural and supernatural reality. At the same time, prospectively Kant redefined the philosophical agenda of the early modern period, determining the problems faced by the nineteenth and twentieth centuries. He changed the very meaning of 'metaphysics' or 'first philosophy' from the first-order study of the supernatural or incorporeal realm of being to the second-order study of the way human inquiry itself makes possible its access to whatever subject matter it studies. He focused attention on the way scientific theories and scientific inquiry are shaped by the creative theorizing of human beings as investigators of nature, and on the way the activity of science relates to other human activities.

Movements as diverse as speculative idealism, logical positivism, phenomenology, and pragmatism all have their foundations in the so-called "Copernican revolution" of Kant's critical philosophy. Kant revolutionized the foundations of philosophical ethics, changing it from a science directed toward achieving a pre-given good, or a study of the way human actions and evaluations are controlled by natural sentiments, into an inquiry into the way free agents govern their own lives according to self-given rational principles.

Kant did all this in part because of the extraordinary breadth of his intellectual curiosity and intellectual sympathies. He first came to the study of philosophy through an interest in the physical sciences: Kant's earliest writings were contributions to the physics, chemistry, astronomy, and geology of his time. Throughout his life, Kant followed developments in the natural sciences: in his late seventies, for instance, he interested himself in Lavoisier's revolution in chemistry, requesting the

crucial experiments to be replicated in Königsberg by a professor of medicine. Kant is generally regarded as the founder of the discipline of physical geography, a subject on which he lectured repeatedly during his university career. An avid reader of travel narratives concerning distant peoples and alien cultures, he reconceptualized the study of anthropology in popularly accessible lectures on the subject delivered over a period of twenty-five years (this was Kant's most frequently given, and most widely attended, university lecture course). As we will see in this book, Kant's philosophical enterprise embraced not only the foundations of scientific knowledge and moral value, but included also revolutionary developments in the history of aesthetics and the philosophy of history. During the last decade of his life Kant also applied his philosophical labors to redefining the relation of reason to religion, and to revolutionizing the theory of international relations by proposing that the permanent relation between states should be one of lawfully ordered peace rather than incipient hostility and eternal preparedness for war.

Kant's achievement is due also to the fact that he so well represents the critical spirit of the eighteenth-century "Enlightenment". This is a spirit of radical questioning and self-reflection that demands of every human activity that it should justify itself before the court of reason. Kant applies this spirit in every area of life: the sciences, aesthetic criticism, morality, politics, above all religion. His position in every field of philosophy is hard to place in the customary categories (such as 'rationalism' and 'empiricism') because it represents both a synthesis of past approaches and a fundamental rethinking of the issues grounding the opposition between traditional schools of thought. In the theory of knowledge Kant is a rationalist, but he limits human knowledge to what can be given in experience. In ethics, he regards human beings as subject to an absolutely binding moral law, but argues that the sole possible authority for such a law is that of their own rational will. In aesthetics, he regards judgments of taste as entirely subjective and non-cognitive, but defends the position that they have a universal validity as strict as that of science or morals. In religion, he regards our own reason as the sole final authority, but denies knowledge to make room for faith.

Like the Enlightenment itself, Kant's philosophy spawned a bewildering variety of thinkers and movements claiming either to be its heirs – or alternatively, or at the same time, to have exposed and corrected its errors. The story of the battle over Kant's legacy and of the struggle to transcend Kant's standpoint amounts to the intellectual history of the entire nineteenth and twentieth centuries. These same conflicts promise to characterize the future in the same way, for as long as we can now foresee and beyond.

The aim of this book is to expound Kant's philosophy. But this first chapter will aim at sketching the life of the man whose philosophy it was.

Background and childhood

Kant was born April 22, 1724 in Königsberg, East Prussia, a seaport located where the River Pregel flows into the Baltic Sea. In Kant's time, the city was an isolated eastern outpost of German culture (though it was occupied by Russian troops for several years during Kant's lifetime). Most of the city was leveled by British and American bombing or by Soviet artillery prior to its invasion by the Soviet army in 1945. After the war it was ethnically cleansed of its German population, renamed Kaliningrad (after a thoroughly hateful Stalinist henchman), and became what it still is, an isolated western outpost of Russian culture. For nearly forty years of the twentieth century, as the headquarters of the Soviet Baltic fleet, it was entirely closed to foreigners and to most Russians as well.

The Lutheran cathedral, where Kant refused on principle to attend religious services, remained a bombed-out ruin until the Gorbachev era, but it was substantially rebuilt and renovated during the 1990s. In Kant's day, the main building of the University (no longer extant) was located near it. The cathedral itself contained the University library where Kant often studied and for a time served as librarian. Kant's tomb, appropriately located *outside* the cathedral on the side (and to the left of the altar), is pockmarked from wartime shrapnel, but remains largely intact (never needing to be rebuilt). It escaped demolition by Allied bombs, reportedly because one Soviet general (with better than average education) ordered that it (together with a statue of Schiller that still stands elsewhere in the city) should be spared the destruction his troops were triumphantly wreaking on the rest of Königsberg. Since the war, the new Russian population of Kaliningrad has kept Kant's tomb constantly adorned with flowers. To this day it is customary for marrying couples to visit it. Apparently the austere rationalist philosopher Immanuel Kant – Lutheran by upbringing but in his maturity always deeply suspicious of popular religious superstition in all its forms – was the nearest imitation of a local Orthodox saint that this old German city had for the new population to venerate.

Eighteenth-century Königsberg was connected to the rest of the world through its access to the sea, and boasted a rich and curiously varied intellectual culture. Nevertheless, it is hardly the place from which one might have expected the greatest revolution in modern philosophy. Nor was Immanuel Kant, judging from his family or his social origins, the sort of person from whom one would have expected such a thing. He was the second son, and the sixth of nine children, born to Johann Georg Kant, a humble saddler (or leather-worker) of very modest means, and Anna Regina Reuter, daughter of a member of the same saddler's guild. Kant believed that his father's family had come from Scotland (and that

the family name had been spelled 'Cant'). He was proud to claim a heritage that would affiliate him with men he admired as much as he did Francis Hutcheson, David Hume, Lord Kames, and Adam Smith. More recent research has shown, however, that he was unfortunately mistaken on this point of his genealogy, probably misled by the fact that more than one of his great uncles had married recent Scottish immigrants. Kant's ancestors, for as far back as they can be traced, were entirely of German stock; his father's family came from Tilsit.

Kant's parents were devout pietists. Pietism was a revivalist movement that arose in the seventeenth century and had a great impact on German culture throughout the eighteenth century. It is comparable to other contemporary religious movements, such as Quakerism or Methodism in England, or Hassidism among central European Jews. Kant's family pastor, Franz Albert Schulz, was also rector of the newly founded *Collegium Fredericianum*. Noticing signs of exceptional intellect in the humble Kant family's second son, he arranged for him an educational opportunity that was surely rare for children of his parents' social class. At the *Fredericianum* Kant was taught Latin and enough else to enter the university at age 16. However, he found the atmosphere of religious zealotry, especially the intellectual tyranny of the catechism, insufferably stifling to both mind and spirit.

In the course of a short treatise on meteorology, he later wrote about the catechisms that "in our childhood we memorized them down to the last hair and believed we understood them, but the older and more reflective we become, the less we understand of them, and on this account we would deserve to be sent back to school once again, if only we could find someone there (besides ourselves) who understood them better" (Ak 8:323).

Attempts are frequently made to identify pietist influences in Kant's moral and religious thought. But virtually all explicit references to pietism in his writings or lectures are openly hostile. He typically identifies pietism either with a spirit of narrow sectarianism in religion or with a self-despising moral lethargy that does nothing to improve oneself or the world but waits passively for divine grace to do everything. Perhaps his mildest remark is one that defines a 'pietist' as someone who "tastelessly makes the idea of religion dominant in all conversation and discourse" (Ak 27:23). Kant's philosophy was in turn regarded with hostility by most of the influential pietists in Königsberg.

It was in the year 1740 that Kant entered the University. The same year Frederick the Great became King of Prussia. This date was significant in the intellectual life of Germany, for one of Frederick's first acts was to recall Christian Wolff from exile in Marburg to his professorship at the University of Halle, thus offering symbolic support to the intellectual movement known as the *Aufklärung* (Enlightenment) of which

Wolff was considered the father. Seventeen years earlier, Wolff had been summarily exiled by Frederick's father, Friedrich Wilhelm I, from Prussian territories under the influence of pietists in the Prussian court. They objected to the way the Enlightenment had made the German universities places of dry scholastic reasoning, rather than religious inspiration and moral exhortation. They also found objectionable Wolff's fascination with "pagan" thought (he was, for instance, one of the first Europeans to undertake the philosophical study of Confucian writings, which he treated in an alarmingly sympathetic spirit). They were also horrified by some of his philosophical doctrines, such as that the human will is subject to causal determination under the principle of sufficient reason (though Wolff did not deny freedom of the will, but was what we would now call a 'compatibilist' or 'soft determinist'). The struggle, both within the universities and in intellectual life generally, between Wolffianism and pietism was decisive for the intellectual environment in which Kant came of age.

Early academic career

The first study Kant took up at the University was Latin literature, which left its mark in the numerous quotations from Latin poets that constitute almost the only literary adornments in Kant's philosophical writings. But soon he came under the influence of those at the university who taught mathematics, metaphysics, and natural science. The best known of these was Martin Knutzen (1713–51), whose early death (it is sometimes speculated) might have deprived him of some of the philosophical influence that was later to be exercised by his most famous student. Knutzen is sometimes described as a Wolffian, but he was more a pietist critic of Wolff than an adherent. Further, it is at best an oversimplification to think of Kant as "Knutzen's student." For one thing, Kant's talents were apparently not much appreciated by Knutzen. He never regarded Kant as among his better students, and this unfortunate fact was largely responsible for what, with hindsight, we now see as the extraordinarily slow development of Kant's academic career. Moreover, Kant's magisterial thesis was completed in 1746 under the direction of Johann Gottfried Teske (1704–72). This makes it more accurate to describe Kant as "Teske's student," though Teske was a natural scientist with few broader philosophical interests. The thesis itself was mainly an elaboration of Teske's researches on combustion and electricity. In fact, all the writings Kant published before the age of 30 were in natural science – on topics in Leibnizian physics, astronomy, geology, and chemistry.

Kant left the University in 1744, at the age of 20, to earn a living as a private tutor, which he did in various households in East Prussia for the next decade. The most influential of his employers was the Count von

Keyserlingk. Even in later years he maintained a social relationship with this family, especially with the Countess. During these years Kant was twice engaged to marry, but both times he postponed marriage on the ground that he was not financially solvent enough to support a family, and both times his fiancée tired of waiting and married someone else. By the time he was financially in a position to marry, he had come to appreciate – probably under the influence of his friend Joseph Green – the independence of a bachelor's life, and had resolved to do without a wife or family.

Kant returned to university life in 1755, receiving the degrees of Master and Doctor of Philosophy, and obtaining a position as *Privatdozent*. This means he was licensed to teach at the University, but was paid no salary, so that he had to earn his living from fees paid him by students for his lectures. Since his livelihood depended on teaching whatever students wanted to learn, he found himself lecturing not only on logic, metaphysics, ethics, natural theology, and the natural sciences – including physics, chemistry, and physical geography – but also on practical subjects that were related to them, such as military fortification and pyrotechnics. For a considerable time Kant devoted his intellectual labors mainly to questions of natural science: mathematical physics, chemistry, astronomy, and the discipline (of which he is now considered the founder) of 'physical geography' – what we would now call 'earth sciences'. This work culminated in *Universal Natural History and Theory of the Heavens* (1755). In this essay Kant was the first to propound the nebular hypothesis of the origin of the solar system. But the financial failure of its publisher had the effect of almost totally suppressing it, and it remained virtually unknown for many years until after La Place had put forward essentially the same hypothesis with greater mathematical elaboration.

In the same year, however, Kant also began to engage in critical philosophical reflections on the foundations of knowledge and the first principles of Wolffian metaphysics, in a Latin treatise *New Elucidation of the First Principles of Metaphysical Cognition*. Here he subjected central propositions and arguments of the Wolffian metaphysics and theory of knowledge to searching criticism, and we find the earliest statement of some of Kant's characteristic thoughts about such topics as causality, mind–body interaction, and the traditional metaphysical proofs for God's existence.

Many years later, in the Preface to his *Prolegomena to Any Future Metaphysics* (1783), Kant made the assertion that it was the recollection of David Hume that first awoke him from his "dogmatic slumbers." There is a literature in German that attempts (rather desperately, in my judgment) to give some sort of biographical substance to this remark.[1] Far more plausibly, Kant's point in making it was to invite his audience

(assumed to have been taught Wolffian philosophy) to find its own path to his critical philosophy through reflection on Hume's skeptical challenges. The juxtaposition of Humean skepticism to Wolffian dogmatism may have been a striking way for Kant to raise the fundamental issue of the possibility of metaphysics, and is certainly indicative of Kant's lifelong admiration for Hume's philosophy. But it is most unfortunate that the remark has been taken as an authoritative autobiographical report about his own philosophical development. For when it is interpreted as saying that Kant began as an orthodox Wolffian metaphysician, only to be roused from complacent rationalism by Hume's skeptical doubts, the remark simply does not correspond at all to the facts of Kant's intellectual life. A student of the development of Kant's philosophy will find that he was from the very start a critic of some of the most basic tenets of Wolffian metaphysics. There never was any "dogmatic slumber" from which to awaken: the long course of Kant's development toward the position of the *Critique of Pure Reason* (and just as significantly, beyond it) was always a restless searching that was terminated only by Kant's eventual decrepitude and death. Its earliest point of departure in 1755 was already a considerable distance from Wolffian 'dogmatism'.

A wider philosophical audience was first attracted to Kant's writings in 1762, when he entered a prize essay competition on the foundations of metaphysics. Moses Mendelssohn won the competition, but Kant's essay, *On the Distinctness of the Principles of Natural Theology and Morals*, won second prize, was published in 1764 along with Mendelssohn's winning essay, and received notable compliments from Mendelssohn (with whom Kant was always on terms of mutual admiration and respect).

Kant's interest in moral philosophy developed relatively late. In the prize essay, as well as his earliest lectures on ethics, he seems to have been attracted by the moral sense theory of Francis Hutcheson. But he was soon to become convinced that a theory based on feelings was inadequate to capture the universal validity and unconditional bindingness of a moral law that must often challenge and overrule corrupt human feelings and desires. His thinking about ethics was dramatically changed about 1762 by his acquaintance with the newly published writings of Jean-Jacques Rousseau: *Émile, Or on Education* and *Of the Social Contract*. Pietism had already taught him to believe in the equality of all human beings as children of God, and in the church universal, encompassing the priesthood of all believers, to be pursued as a moral ideal in a sinful world of spiritual division and unjust inequality. These convictions now took the more rationalistic form of Rousseau's vision of human beings, free and equal by nature, who find themselves in an unfree social world where the poor and weak are oppressed by the rich and powerful. Soon Kant began defining his own ethical position

through emphasis on the sovereignty of reason, associating his moral philosophy with the title 'metaphysics of morals'. However, it was another twenty years before Kant brought his ethical theory to maturity. In the meantime, the task to which he devoted his principal labor was that of reforming the foundations of the sciences and discovering the proper relation within them between empirical science and the claims of *a priori* or metaphysical knowledge.

Kant's closest friend during his youth was Johann Daniel Funk (1721–64), a professor of law, who led a rather wild life and died at an early age. Like his friend Funk (and contrary to the grossly distorted traditional image of him), Kant was always a gregarious man, thought of by those who knew him as charming, witty, and even gallant. Compared to Funk, however, he was also much more self-controlled and prudent. His sociability included regular play at cards and billiards, which he did with notable shrewdness and skill. Kant's winnings often supplemented his meager academic income. After Funk's death, Kant made his longest and most intimate friendship, with the English businessman Joseph Green (1727–86). Green was an eccentric bachelor and a man of very strict and regular habits. It is probably through Green's influence that Kant acquired many of the characteristics pertaining to the (often highly distorted) picture that was later formed of him. From quite early on, Kant invested his savings in the mercantile ventures of the firm of Green & Motherby, which was profitable enough to provide Kant with a comfortable fortune by the time he gained his professorship in 1770.

Genesis of the critical philosophy

The slow development of Kant's academic career corresponds to the long gestation period of the system of thoughts for which we now most remember him. Professorships in logic and metaphysics became open at the University of Königsberg in 1756 and 1758, but Kant did not even apply for the first, and with his still very limited qualifications he was routinely passed over for the second. After the recognition he received from Mendelssohn and the Prussian academy, he was offered a professor of poetry at the university in 1764, but declined it because he wanted to continue devoting himself to natural science and philosophy. In 1766 he did accept a position as sublibrarian at the University, providing him with his first regular academic salary. But he declined opportunities for professorships in 1769, first at Erlangen and then at Jena, chiefly because of his reluctance to leave East Prussia, but also because he expected the professorship of logic at Königsberg would be available to him the following year. In subsequent years he had other opportunities (for instance, he was offered a professorship at Halle in 1778), but chose never to leave Königsberg. Just as Beethoven, the most revolutionary of

all composers, wrote some of his most original music after he was totally deaf, so Kant, the most cosmopolitan of all philosophers, lived in an isolated province of northeastern Europe and never traveled farther than thirty miles from the place of his birth.

In the Latin inaugural dissertation he wrote on assuming his professorship at Königsberg, *On the Forms and Principles of the Sensible and Intelligible World*, Kant took several important steps in the direction we can now see eventually led him to the 'critical philosophy' of the 1780s and 1790s. By 1772, Kant told his friend and former student Marcus Herz that he was at work on a major philosophical treatise, to be entitled *The Limits of Sensibility and Reason*, which he expected to finish within a year. But it was nearly a decade more before Kant published the *Critique of Pure Reason*. During the 1770s Kant wrote and published very little. Despite his elevation to a professorship, Kant continued to live in furnished rooms on the island in the Pregel on which stood both the University building and the cathedral in which its library was housed. It would be another thirteen years before he was able to purchase a house of his own.

Early in this "silent decade," however, Kant began lecturing on the subject of 'anthropology', stimulated (or provoked) by Ernst Platner's *Anthropology for Physicians and Philosophers* (1772). Kant rejected Platner's 'physiological' reductivism in favor of an approach that emphasized the practical experience of human interaction and the historicity of human beings. Yet Kant was always deeply skeptical of the capacity of human beings to gain anything like a scientific knowledge of their own nature, and he was especially dissatisfied with the entire state of the study of human nature up to now, looking forward to a future scientific revolution in this area of study (which he himself did not pretend to be able to accomplish). He lectured on anthropology in a popular style for the next twenty-five years. These lectures were the most frequently given and the most well attended of any he gave during his teaching career. Kant's ideas about anthropology exercised a powerful but subtle influence on his treatment of epistemology, philosophy of mind, ethics, aesthetics, and the philosophy of history, but it is an influence difficult to assess because Kant never articulated a systematic theory of anthropology, and his published writing on anthropology was limited to a popular textbook derived from his lectures, *Anthropology from a Pragmatic Standpoint* (1798) which he issued at the end of his teaching career.

The *Critique of Pure Reason* was finally published in the spring of 1781 (less than a month before Kant's 57th birthday). Although Kant brought his labors on it to a conclusion very rapidly, in the space of about four months in 1779–80, this book had been nearly ten years in preparation. It is reported that he had read every sentence of it to Green, whose opinion even in philosophical matters he valued very highly. Once the

Critique was published, the evident originality of the thoughts contained in it and the difficulty of his struggle to achieve them both led Kant to expect that it would attract immediate attention at least among philosophers. He was therefore disappointed by the cool and uncomprehending reception it initially received. For the first year or two he received from those whom he most expected to give his book a sympathetic hearing only a bewildered silence.

Kant found especially frustrating the review of the *Critique* published in the *Göttingen Learned Notices* in January, 1782. It was ostensibly written by Christian Garve (a man Kant respected) but had been heavily revised by the journal's editor, J. G. Feder, a popular Enlightenment philosopher of Lockean sympathies who had little patience for metaphysics in any form and no sympathy at all for the abstruse project in which Kant was engaged. The review interpreted Kant's transcendental idealism as no more than a variation on Berkeley's idealism – a reduction of the real world to subjective representations, based on an elementary confusion between mental states and their objects. The review, together with the evident incomprehension of the *Critique* by most of its earliest readers, caused him to attempt a more accessible presentation of his ideas in *Prolegomena to Any Future Metaphysics* (1783). But Kant was not a good popularizer, and it would be several more years before the *Critique* began to get the kind of attention Kant had hoped for.

Years of academic success

Kant was born poor, and he remained poor, an unsalaried, marginal academic, well into middle age. But his investments with Green and his appointment to a professorship finally gave him a comfortable living. And by the early 1790s his fame made him one of the highest paid professors in the Prussian educational system. During the late 1760s and for most of the 1770s he lived, along with many others from the University, in a large rooming house owned by the publisher and bookdealer Kanter. In 1783, at age 59, through the help and influence of his friend, the mayor of Königsberg, Theodor Gottlieb von Hippel (1741–96), Kant finally bought a home of his own – a large, comfortable house on Prinzessinstraße in the center of town, almost in the shadow of the royal castle that gave the city its name.

Kant's friend Hippel was a remarkable man. He was not only active politically, but also intellectually. He was a learned and intelligent man, the author of whimsical, satirical plays and novels in the style of Sterne. He also wrote progressive political treatises defending the civil equality of Jews, and argued for a quite radical position on the social status of women, advocating the reform of marriage to insure their equality with men in all spheres of life. Hippel's views on the emancipation of women

were far in advance of Kant's own, even though at the time rumor had it that Kant shared in the authorship of these "feminist" writings. If Kant in later years was comfortably off, Hippel was downright rich. Kant had come to know Hippel in the same circle as Funk, and Hippel's lifestyle had been more influenced by Funk as well. After Hippel's death (and like other male advocates of women's rights in that age, such as William Godwin), he was the object of disapproving rumors regarding his scandalous sexual behavior. Kant, however, always refused to participate in these attacks.

Another of Kant's notable friendships is even more curious – the one with J. G. Hamann (who was also a close friend of Green). Hamann was a thinker and writer of great brilliance, but his views – like his personality – could hardly have been more different from Kant's. Hamann was an eccentric religious thinker, who combined philosophical skepticism with fideist irrationalism. He had a troubled life-history, lived an unconventional life (for instance, cohabiting with a woman he never married), and was an imprudent, unstable, unhealthy man. Hamann's writings are terse, impressively learned, full of idiosyncrasies, ironies, and inventive allusions, always tantalizingly (or infuriatingly) cryptic. He was a trenchant critic of the Enlightenment, including Kant's philosophy, and a mentor of both the German counter-enlightenment and the *Sturm und Drang* literary movement. It says something very significant, and very favorable, about both men's characters and the largeness of both their minds, that they were genuinely friends, and that their profound differences in style and outlook apparently never led to any significant personal estrangement.

Kant's relation with other friends and acquaintances reveals a more ambiguous picture. During the 1760s he was close to the customs official Johann Konrad Jacobi and perhaps even more so to his wife Maria Charlotta.[2] But when she left her husband and took up with another acquaintance of Kant, master of the mint Johann Julius Göschel, after the divorce and remarriage Kant broke off relations with the adulteress and refused ever to see her or her new husband. He was not always so intolerant of sexual indiscretions, however. When his doctoral student F. V. L. Plessing[3] fathered an illegitimate child in 1784, Kant undertook the responsibility of conveying the necessary payments to the young woman, and may even have supplied some of the funds himself. Yet when in 1794 a troubled young woman, Maria von Herbert, sought the philosopher's advice and consolation in a time of inner anguish and despair, Kant showed remarkable insensitivity to her feelings, dismissing her to their mutual friend Elizabeth Motherby as "die kleine Schwärmerin" (the little enthusiast), and citing her as a sad example of what can happen to young women who do not control their fantasies. Some years later, Maria committed suicide.

Students whom Kant regarded as straying from the proper path were sometimes dealt with unkindly. When Kant's former student J. G. Herder criticized Kant in the first two volumes of his *Ideas for the Philosophy of History of Humanity* (1785–7), Kant wrote somewhat condescending reviews of Herder's work, then tried to pass the dubious task of criticizing him along to another of his highly able students, Christian Jacob Kraus (who was the chief exponent of Adam Smith's economic theories in Germany). When Kraus refused to comply with Kant's wishes, they quarreled and their previously close friendship came to an end. Kant helped the young J. G. Fichte to begin his philosophical career by aiding him in the publication of his first work, *Attempt at a Critique of All Revelation* (1792). But in 1799, perhaps under the jealous influence of some of his students, Kant publicly denounced Fichte, disclaiming him as a follower of the critical philosophy and citing the Italian proverb: "May God protect us from our friends, for we shall manage to watch out for our enemies ourselves" (Ak 12:371).

Kant's house on Prinzessinstraße

The first floor of Kant's house contained a hall in which he gave his lectures, and the kitchen where food was prepared by a female cook (he could now finally afford to hire one); on the second floor was a sitting room, a dining room, and Kant's study (where there reportedly hung over his writing desk the only decoration he permitted in the house – a portrait of Rousseau). Kant's bedroom was on the third floor. For many years Kant had a personal servant, Lampe – who, however, was apparently given to drink, and was discharged in the late 1790s when he reportedly attacked his frail and aging master during a quarrel.

In the second-floor dining room Kant enjoyed his only real meal of the day, a dinner at which he usually entertained several guests. Königsberg was a seaport, and although Kant never himself ventured far from it, he took the opportunity to acquaint himself with many of the distinguished foreigners who passed through. By the time of these banquets (in the early afternoon), Kant had usually completed his main academic work. He rose regularly at 5 a.m., having only a cup of tea and a pipe of tobacco for breakfast. Then he prepared for his lectures, which he delivered five or six days a week, beginning at 7 or 8 in the morning. After them, he would go his study and write until time for dinner. After his guests had departed, Kant would often take a nap in an easy chair in his sitting room (sometimes a good friend, such as Green, would nap in the chair next to him). At 5 p.m. the philosopher would take his constitutional walk, whose timing, according to the famous legend, was so precise and unvarying that the housewives of Königsberg could set their clocks by the minute at which Professor Kant walked past their windows. Yet the

regularity of Kant's schedule, as well as his crochets about his health and especially his diet (he believed in eating a lot of carrots, and drank wine daily, but never beer), probably resulted less from a compulsive personality than from the necessity of an aging man, who had never been in the best of health, to keep himself strong enough to complete philosophical labors which he had not been able properly to begin until he was far into middle age. Kant's evenings were often spent socializing, either at Green's house, or Hippel's, or with the Count and Countess Keyserlingk.

Enlightenment and philosophy of history

In the middle of the 1780s, Kant laid the foundation for much of nineteenth-century philosophy of history in several brief occasional essays. To a significant degree, Kant's thinking about history was prompted by his reading of Herder's *Ideas*. Herder saw himself as a critic of the Enlightenment rationalism Kant defended, and Kant's contributions to the philosophy of history were in part an attempt to vindicate the cause of Enlightenment in that debate. In 1786 Kant added to these reviews a satirical essay, *Conjectural Beginning of Human History*, parodying Herder's use of the Genesis scriptures in Book 10 of the *Ideas* to support his anti-Enlightenment theory of human history. But the *Conjectural Beginning* also makes some serious points both about the use of imaginative conjectures in devising such narratives and about the role of reason and conflict in the progressive historical development of humanity's faculties.

Another important short essay displaying the historical conception of Kant's philosophy was prompted by the published remark of a conservative cleric, who dismissed the call for greater enlightenment in religious and political matters with the comment that no one had yet been able to say what was meant by the term 'enlightenment'. Kant's response was the short essay *An Answer to the Question: What is Enlightenment?* (1784). Kant refuses to identify *enlightenment* with mere learning or the acquisition of knowledge (which he thinks is at most a consequence of that to which the term genuinely refers). Instead, Kant regards enlightenment as the act of leaving behind a condition of immaturity, in which a person's intelligence must be guided by another. Many people who are able to direct their own understandings, or would be able if they tried, nevertheless prefer to let others guide them, either because it is easy and comforting to live according to an established system of values and beliefs, or because they are anxious over the uncertainties they will bring upon themselves if they begin to question received beliefs or afraid of taking on the responsibility for governing their own lives. To be enlightened is therefore to have the courage and resolve to be self-directing in one's thinking, to *think for oneself*.

Kant also emphasizes that enlightenment must be regarded as a social and historical process. Throughout humanity's past, most people have been accustomed to having their thinking directed by others (by paternalistic governments; by the authority of old books; most of all, and most degrading of all, in Kant's view, by the priestcraft of religious authorities who usurp the role of individual conscience). Becoming enlightened is virtually impossible for an isolated individual, but it becomes possible when the practice of thinking critically becomes prevalent in an entire public in which reigns a spirit of free and open communication between its members. Kant's proposals concerning freedom of communication in *What is Enlightenment?* are based not on any alleged individual right to freedom of expression, but are entirely consequentialist in their rationale and tailored to his time and place, designed to encourage the growth of an enlightened public under the historical circumstances in which he found himself.

One unjust calumny often directed against the Enlightenment is that it was a movement devoid of a sense of the historical or an awareness of the historical context of human actions and endeavors. The charge is perniciously false, and especially so when directed toward Kant. What it often represents is a deceptive presentation of a different view of history from the Enlightenment's, or else an even shabbier attempt by nineteenth-century thinkers to pass off the Enlightenment's accomplishments in historical thinking as their own, or both of these at once. The *Critique of Pure Reason* (even its title) reflects a historical conception of Kant's task. Kant sees the 'critique' as a metaphorical court before which the traditional claims of metaphysics are being brought to test their validity. His metaphor is drawn from the Enlightenment political idea that the traditional claims of monarchs and religious authorities must be brought before the bar of reason and nature, and henceforth the legitimacy of both should rest only on what reason freely recognizes. Kant's philosophy is self-consciously created for an age of enlightenment, in which individuals are beginning to think for themselves and all matters of common interest are to be decided by an enlightened public through free communication of thoughts and arguments.

For nearly twenty years, Kant had intended to develop a system of moral philosophy under the title 'metaphysics of morals'. It is probably no accident that he began to fulfill this intention only after he had been provoked into thinking about human history and the moral predicament in which the natural progress of the human species places its individual members. The *Groundwork for the Metaphysics of Morals* (1785) is one of the classic works in the history of ethics, and (as its title implies) it proposes to lay the ground for Kant's ethical system. But it never claims to do more than provide the fundamental principle of the system. It

discusses the application of the moral principle only by way of selected illustrations, and does not provide us with a systematic theory of duties. During the next decade, Kant continued to reflect both on the foundations of ethics and on the application of his ethical principles to morality and politics. But he presented something like an ethical system only at the very end of his career, in the *Metaphysics of Morals* (1797–8). Kant's ethical thought, and even what is said in the *Groundwork* itself, is often misunderstood because these later works are not taken into account in reading it.

In 1786 Kant's philosophy was suddenly thrust into prominence by the favorable discussion of it presented in a series of articles in Christoph Wieland's widely read publication *Teutsche Merkur* (called "Letters on the Kantian Philosophy") by the Jena philosopher Karl Leonard Reinhold. Reinhold's presentations of Kant did very suddenly what Kant's own works had thus far failed to do – namely, to make the theories of the *Critique* into the principal focus of philosophical discussion in Germany. Soon the critical philosophy came to be seen as a revolutionary new standpoint; the main philosophical questions to be answered were whether one should adopt the Kantian position, and if one did, exactly what version or interpretation of it one should adopt. Soon there also arose a new kind of *critic* of Kant's philosophy – an irrevocably "post-Kantian" philosopher, whose criticisms were motivated by alleged unclarities and tensions within Kant's philosophy itself. These critics sought to absorb the lessons of the Kantian philosophy and yet also to "go beyond" it.

For this reason, and because of the misunderstandings to which Kant had discovered his position was subject, he decided to produce a second edition of the *Critique*, in which he could present his position more clearly. At first he thought he would add a section on *practical* (or moral) reason, following up his treatment in the *Groundwork* (and also replying to critical discussions of that work that had appeared). In 1787 the new and improved version of the *Critique of Pure Reason* did appear, but by then Kant had decided that his discussion of practical reason would have to be too lengthy to be added to what was already a very long book, so he decided to publish it separately as a second 'critique'.

Within a short time, Kant was working on a third project that was to bear a parallel title. Kant conceived of philosophy as an architectonic system, but it was never part of his systematic project to write three 'critiques'. The *Critique of Practical Reason* grew opportunistically out of Kant's desire to respond to critics of his *Groundwork*, and also from his decision to revise the *Critique of Pure Reason* – as stated above, he originally intended to include a "critique of practical reason" in this second edition, but wrote a separate book when he saw that the length of this new section was getting out of hand. Kant's reasons for writing the

Critique of the Power of Judgment were complex, and a bit inscrutable, as is the work itself. Kant had been thinking for a long time about the topic of taste and judgments of taste, and wanted to come to terms with the modern tradition of thinking about these matters, found in such philosophers as Hutcheson, Baumgarten, Hume, and Mendelssohn. Judgments of taste, such as that something is beautiful or ugly, have the peculiarity that on the one hand they do not ascribe a determinate objective property to an object but report merely the subject's own pleasure or displeasure in it, and yet on the other hand they do claim a kind of quasi-objectivity, as though there are some things which *ought* to please or displease all subjects. Kant was dissatisfied with both Baumgarten's attempt to analyze beauty as perfection experienced by the senses rather than by the intellect and by Hume's view that taste is merely pleasure or displeasure in an object considered in relation to certain normative conditions of experiencing it, such as disinterestedness. He wanted to understand how the workings of our cognitive faculties themselves, especially the harmony between sensible imagination and understanding required for all cognition, might play a role in generating an experience that was at once subjective and yet normative for all. But to solve this problem is far from being the whole motivation behind the third *Critique*.

The two main themes dealt with in this work – aesthetic experience and natural teleology – were both preoccupations of the Enlightenment's critics, such as Herder. Kant also needed to clarify and explicate his own thinking about the status of teleological thinking in relation to natural science, a subject that had engaged him before in essays about natural theology and the philosophy of history. But if we are to take him at his word, the main motive for writing the *Critique of the Power of Judgment* was to deal with the "immense gulf" that he saw between the theoretical use of reason in knowledge of the natural world and its practical use in morality and moral faith in God. It remains to this day a subject of controversy exactly how Kant hoped to bridge this gulf in the third *Critique* and how far he was successful. But the *Critique of the Power of Judgment* reveals Kant, now in his late sixties, as a philosopher who is still willing to question and even revise the fundamental tenets of his system. And to his idealist followers, Fichte, Schelling, and Hegel, it was the *Critique of the Power of Judgment* that seemed to them to show Kant as open to the kind of radical speculative philosophy in which they were interested.

A decade of struggle and decline

The final decade of Kant's activity as a philosopher was one beset with conflict, and well before the end of it, Kant's health and even his mental powers were very much in decline. As the critical philosophy became

increasingly prominent in German intellectual life, and as it came to be variously interpreted by different proponents and would-be reformers of it, Kant found himself defending his position on several sides, against the attacks of Wolffians such as J. A. Eberhard, Lockeans such as J. G. Feder and C. G. Selle, popular Enlightenment rationalists such as Christian Garve, religious fideists such as Thomas Wizenmann and F. H. Jacobi, or against a new kind of 'Kantian' speculative philosopher, such as the brilliant but difficult Salomon Maimon. Kant's larger-scale published works during the 1790s, however, were devoted to applying the critical philosophy to matters of general human concern, especially in the practical sphere – to religion, political philosophy, and to the completion of the ethical system he had for thirty years called the 'metaphysics of morals'.

Kant also came into conflict with the political authorities over his views on religion. From the beginning of Kant's academic career until 1786, the Prussian monarch had been Frederick the Great. Frederick may have been a military despot, but his views in matters of religion favored toleration and theological liberalism. Many considered him to be privately a "freethinker" or even an outright atheist. Frederick's death in 1786 brought to the throne a very different sort of monarch, his nephew Friedrich Wilhelm II, for whom religion was a very serious matter. The new king had long been shocked by the wide variety of unorthodoxy, skepticism, and irreligion that had been permitted under his uncle to flourish within the Prussian state and even within the Lutheran state church. Two years after coming to power, he removed Baron von Zedlitz (the man to whom Kant had dedicated the *Critique of Pure Reason*) from the position of Minister of Education, replacing him with J. C. Wöllner (whom Frederick the Great had described as a "deceitful, scheming parson"). Both the king and his new minister believed that the stability of the state depends directly on correct religious belief among its subjects, and hence that those who questioned Christian orthodoxy were directly threatening the foundations of civil peace. To them, Kant's attack on objective proofs for God's existence, and his denial of knowledge to make room for faith, seemed dangerously subversive. And his enlightenment principles – that all individuals have not only a right but even a duty to think for themselves in religious matters, and that the state should encourage such free thought by protecting a "public" realm of discourse from all state interference – these seemed to the new king and his orthodox followers like recipes for civil anarchy.

Wöllner soon issued two "religious edicts" intended to reverse the effects of Enlightenment thinking on both the church and the universities, by subjecting clergy and academics to tests of religious orthodoxy concerning both what they published and what they taught from the pulpit or the lectern. The edicts put many liberal pastors in the position

of choosing between losing their livelihood and teaching what they regarded as a set of outdated superstitions. Action was taken against some academics as well (especially critical biblical scholars), who were forced either to recant what they had said in their writings (which usually discredited them among their colleagues) or else to lose their university positions (and with them any opportunity to teach their views at all). Writings on religious topics were also to be submitted to a board of censorship, which had to approve the orthodoxy of what they taught before they could be published.

By 1791 Kant learned from his former student J. G. Kiesewetter, who was a royal tutor in Berlin, that the decision had been taken to forbid him to write anything further on religious subjects. But by this time Kant's prominence was such that this would not be an easy or a comfortable action for the reactionary ministers to take. Kant had planned to write a book on religion, and did not let word of these threats dissuade him. But he very much wanted to avoid confrontation with the authorities, both in order to protect himself and on sincerely held moral grounds.

Kant was far from being a political radical on matters such as this. His political thought is strongly influenced by the Hobbesian view that the state is needed to protect both individuals and the basic institutions of society against the human tendencies to violent infringement of rights, and that in order to prevent civil disorder, the state must have considerable power to regulate the lives of individuals. *What is Enlightenment?* teaches that it is entirely legitimate for freedom of communication to be regulated in matters that are "private," dealing with a person's professional responsibilities. This principle might have been used to justify the very actions that had been taken by the Prussian government against pastors and even professors, insofar as their unorthodox teachings were expressed in the course of discharging their clerical or academic duties. He deplored Wöllner's edicts, of course, and regarded their application to the clergy only as having the effect of making hypocrisy a necessary qualification for ecclesiastical office. But it is not at all clear whether he regarded these measures as anything worse than disastrously unwise abuses of the state's legitimate powers. Kant sincerely believed that it is morally wrong to disobey even the unjust commands of a legitimate authority, unless we are commanded to do something that is in itself wrong. Even before anything was done to him he had made the decision that he would comply with whatever commands were made of him. This is all quite clear in Kant's first extensive presentation of his philosophy of the state in the second part of the three-part essay he wrote on the common saying: "That may be correct in theory but it does not work in practice." There he defends (against Hobbes) the position that the subjects of a state have some rights against the state which are binding on

the government but not enforceable against the head of state. This means that there can be no right of insurrection, and that even the unjust commands of a legitimate authority must be obeyed by its subjects (so long as these do not directly command the subject to do something that is in itself wrong or evil). The application of this last principle to Kant's own situation is obvious: He had decided that when the Prussian authorities command him to cease writing or teaching on religious subjects, he would obey them.

But of course Kant had no intention of anticipating such commands, or doing anything merely to please authorities he regarded as unenlightened, unwise, and unjust. And he was determined to make use of all the legal devices at his disposal to thwart their intentions. In 1792, when he gave his essay on radical evil (which later became Part One of the *Religion Within the Boundaries of Mere Reason*) to the *Berlin Monthly* for publication, he insisted on its being submitted to the censorship; when it was rejected, he submitted the entirety of the *Religion* to the academic faculty of philosophy in Jena, which under the law was an alternative to the official state censorship. A first edition appeared in 1793, and a second (expanded) edition in 1794. Kant's evasion angered the censors in Berlin, however, and led them finally to take the action against him they had been planning. In October, Wöllner sent Kant a letter expressing in the King's name the royal displeasure with his writings on religion, in which "you misuse your philosophy to distort and disparage many of the cardinal and basic teachings of the Holy Scriptures and of Christinaity" (Ak 7:6). It commanded him neither to teach nor write on religious subjects until he was able to conform his opinions to the tenets of Christian orthodoxy. In his reply, Kant defended both his opinions and the legitimacy of his writing about them, but did solemnly promise to the King that he would obey the royal command (Ak 7:7–10).

Even the title of the *Religion* was carefully crafted by Kant in light of what he took the legal situation to be. Kant regarded revealed theology (based on the authority of the Church and scripture) as a "private" province of those whose profession obliges them to accept that authority. But when an author writes on religion apart from appeal to such authorities, basing his assertions solely on reason unaided by any appeal to revelation, he is writing for the 'public' sphere. In fact, Kant's *Religion* is an attempt to provide an interpretation, in terms of rational morality, of central parts of the Christian message – original sin, salvation through faith in Christ, the vocation of the Church. Its principal aim is to convince Christians that their own religious beliefs and experience are entirely suitable vehicles for expressing the moral life as an enlightened rationalist philosopher understands it. No doubt Kant's rationalistic reinterpretations were (and still are) apt to seem abstract and bloodless to many Christians. There is no role in Kant's account of salvation for

vicarious atonement made by the historical person of Jesus Christ. His rational religious faith has no room for miracles, disapproves of religious practices such as petitionary prayer, and Kant regards religious rites as "superstitious pseudo-service of God" when they are presented as necessary for moral uprightness or justification of the sinner before God. He directly attacks the *Pfaffentum* ("priestcraft" or "clericalism") of a professional priesthood, looking forward to the day when the degrading distinction between clergy and laity will disappear from a more enlightened church than now exists. (As I have already mentioned, Kant's own conduct reflected his principles. He refused on principle to participate in religious liturgies. Even when his ceremonial position as rector of the University of Königsberg required him to attend religious services, he always declined, reporting that he was "indisposed".)

The *Religion* has much to tell students of Kant's ethical theory both about its moral psychology and about the application of moral principles to human life. The essay on radical evil makes it clear that for Kant moral evil does not consist merely in determination of the will by natural causes (as it may sometimes seem to do from what is said in the *Groundwork* or even the second *Critique*). Instead, the essay on radical evil insists that all moral choice consists in the adoption of a maxim (whether good or evil) by a free power of choice, and thus transcends the natural causality Kant takes to be incompatible with freedom. It also coheres with Kant's philosophy of history in presenting the social condition, and the natural propensity to competitiveness awakened in it, as the ground of all moral evil. Part Three of the *Religion* argues that since the source of evil is social the moral progress of individuals cannot come from their isolated strivings for inner purity of will but can result only from their freely uniting themselves in the adoption of common ends. The ideal "realm of ends" is therefore to receive earthly reality in the form of a "people of God" under moral laws, who are to unite *freely* (not in the form of a coercive state) and *universally* (not as an ecclesiastical organization limited by creeds and scriptural traditions). The essence of religion for Kant consists in recognizing the duties of rational morality as commanded by God, and in joining with others to promote collectively the highest good for the world. It is in this free form of religious association, and not the coercive political state, that Kant ultimately places his hopes for the moral improvement of the human species in human history. The role of the state in history for Kant is not to provide the human species with its final aim, but rather to provide the necessary conditions of external freedom and justice in which the moral faculties of human beings may develop, and free (religious) forms of association may flourish in peace.

Kant had been forbidden by the authorities to write on religious topics, but he had no intention of keeping quiet on other matters of general

human concern, even when his views were likely to be unpopular with the government. In March, 1795 a period of war between the revolutionary French Republic and the First Coalition of monarchical states was brought to a close by the Peace of Basel between France and Prussia. Kant's essay *Toward Perpetual Peace* should be read as an expression of support not only for this treaty but also directly for the First French Republic itself, since here he declares that the constitution of every state should be republican and also conjectures that peace between nations might be furthered if one enlightened nation transformed itself into a republic and then through treaties became the focal point for a federal union between other states. Kant begins with four "preliminary articles" designed at promoting peace between nations through their conduct of themselves under the present condition of incipient warfare and the diplomatic conduct surrounding it. The essay then proceeds to three "definitive articles" defining a relationship between states that will lead to a condition of peace that is not merely a provisional and temporary interruption of the perpetual condition of war but constitutes a permanent or "eternal" condition of international peace. This is followed by two "additions" outlining the larger philosophical (historical and ethical) presuppositions of Kant's approach, and an appendix in which Kant discusses the manner in which politicians or rulers must conduct affairs of state if they are to be in conformity with rational principles of morality.

Toward Perpetual Peace is the chief statement authored by a major figure in the history of philosophy that addresses the issues of war, peace, and international relations that have been central concerns of humanity during the two centuries since it was written. Kant drew his inspiration from the *Project for Rendering Peace Perpetual in Europe* by the Abbé de Saint-Pierre (1712), and comments on it by Jean-Jacques Rousseau (1761). But his aims in *Toward Perpetual Peace* are much more ambitious in that their scope is not limited to the Christian nations of Europe but motivated by universal moral principles. His purpose is not merely to prevent the destruction and bloodshed of war, but even more to effect peace with justice between nations as an indispensable step toward the progressive development of human faculties in history, in accordance with the philosophy of history he projected over a decade earlier. *Toward Perpetual Peace* is perhaps Kant's most genuine attempt to address a universal enlightened public concerning issues of importance not only to scientists and philosophers but vital to all humanity.

The history of Kant's conflict with, and for a time his submission to, the Prussian authorities regarding religion, has an unexpectedly happy ending. Friedrich Wilhelm II, typical of rulers in all ages who make a display of religious orthodoxy central to their conception of public life,

permitted himself a private lifestyle that was morally unconventional, and the reverse of prudent, temperate, or healthy. When he died rather suddenly in 1797, Kant chose (in a spirit more wily than submissive) to interpret his earlier promise to abstain from writing on religion as a personal commitment to this individual monarch, and regarded the latter's death as freeing him from the obligation. The royal censors, who were always regarded within the hierarchy of the Lutheran church as uncultured fanatics, probably never had the power to enforce their prohibitions against Kant anyway, and certainly lacked it once the king was dead. In the *Conflict of the Faculties* (1798) Kant had his final say on religious topics, framing his discussion in terms of an account of academic freedom within the state that vindicated his course of action in publishing the *Religion* several years earlier (the act that had provoked the royal reproof). As for Kant's persecutor Wöllner, who had risen to the nobility from a rather lowly background on the strength of his devotion to the cause of religious conservatism, he had already been treated with conspicuous ingratitude by the fickle king whose religious prejudices he had done his best to serve. Soon after the death of Friedrich Wilhelm II, he lost whatever influence he ever had over Prussian educational and ecclesiastical policies, and eventually died in poverty.

Old age and death

Kant retired from university lecturing in 1796. He then devoted himself to three principal tasks. The first was the completion of his system of ethics, the *Metaphysics of Morals*, consisting of a Doctrine of Right (covering philosophy of law and the state) and a Doctrine of Virtue (dealing with the system of ethical duties of individuals). The first part was published in 1797 and the whole in 1798. Kant's second task was the publication of materials from the lectures he had given over many years. He himself published a text based on his popular lectures on anthropology in 1798. Declining powers led him to consign to others the task of publishing his lectures on logic, pedagogy, and physical geography that appeared during his lifetime.

Kant's third project after his retirement is the most extraordinary. He set out to write a new work centering on the transition between transcendental philosophy and empirical science. In it Kant was responding creatively both to recent developments in the sciences themselves (such as the revolution in chemistry initiated by Lavoisier's investigation of combustion) and to the work of younger philosophers who took their inspiration from the Kantian philosophy itself (such as the "philosophy of nature" of F. W. J. Schelling, who was still in his early twenties). Kant's failing powers prevented him from completing this work, but from the fragments he produced (that were first published in the early

twentieth century under the title *Opus Postumum*), we can see that even in his late seventies, Kant still took a critical attitude toward every philosophical question and especially toward his own thoughts. Even while struggling against the failure of his intellectual powers, he was also fighting to revise in fundamental ways the critical philosophical system whose construction had been the labor of his entire life. In this way, the next generation of German philosophers, who saw it as their task to "go beyond Kant", were thinking more fundamentally in Kant's own spirit than have been the generations of devoted Kantians since, who ever and again want to go "back to Kant" and who tirelessly attempt to defend the letter of the Kantian texts against the attempts of his first followers to extend and correct his philosophy. Kant died February 12, 1804, a month and a half short of his eightieth birthday.

further reading

Lewis White Beck, *Early German Philosophy*. Cambridge, MA: Harvard University Press, 1969.

—, *Mr. Boswell Dines with Professor Kant*. Bristol, England: Thoemmes Press, 1995.

Frederick Beiser, *The Fate of Reason*. Cambridge, MA: Harvard University Press, 1987.

Ernst Cassirer, *Kant's Life and Thought*, tr. James Haden. New Haven: Yale University Press, 1981.

Eckart Förster, *Kant's Final Synthesis*. Cambridge, MA: Harvard University Press, 2000.

Manfred Kuehn, *Kant: A Biography*. Cambridge, England: Cambridge University Press, 2001.

Paul Guyer (ed.), *The Cambridge Companion to Kant*. New York: Cambridge University Press, 1992.

notes

1 For instance, see Hans Gawlick and Lothar Kriemendahl, *Hume in der deutchen Aufklärung: Umrisse der Rezeptionsgeschichte* (Stuttgart: Frommann-Holzboog, 1987).

2 One of Maria Charlotta's extant letters to Kant reads: "I lay claim to your society tomorrow afternoon. 'Yes, yes I will be there,' I hear you say. Good, then, I will expect you, and then my clock will be wound as well" (Ak 10:39). Much is read into this last figure of speech by a few Kant scholars who apparently want to entertain the desperate hope that Kant may not after all have been a lifelong celibate.

3 The troubled, romantic Plessing was also an acquaintance of Goethe, and is the subject of his poem "Harzreise im Winter," which later provided the text of Brahms' *Alto Rhapsody*, op. 53.

chapter 2

synthetic *a priori* cognition

I the possibility of metaphysics

One of the chief tasks of the philosophers we most remember from the early modern period was to provide a general philosophical foundation for the emerging physical science – both of its image of nature and of its method of inquiring into nature – while coming to terms in the process with the traditional metaphysical and religious picture of the world. This was the chief focus of Descartes' philosophical endeavors, and this is why he is regarded as the founder of modern philosophy. Most other modern philosophers were his followers in this respect, making one kind of modification or another in this fundamental project. Kant's *Critique of Pure Reason* is the last great attempt to achieve the same aims. At the same time, it inaugurated a new way of achieving them that would result in a radical transformation of the aims themselves. Kant conceived the foundation of natural science, as well as the rational attempt to address the most universal concerns of humanity, as the province of a special and fundamental science – the science of metaphysics. But the way he chose to do metaphysics opened up an entirely new approach to philosophy, which he dubbed 'transcendental'.

Kant understood the term 'meta-physics' (etymologically, 'beyond nature') in epistemological terms. That is, for the purposes of metaphysics, 'nature' is what is known through experience, and so 'metaphysics' is a science demarcated not by the set of objects with which it deals but by the *a priori epistemic status* of its principles. Kant also posed his philosophical task regarding metaphysics in historical terms, as the need to address a crisis of legitimacy in metaphysics, and thus to establish metaphysics on a basis that might henceforward ensure its legitimacy. In the Preface to the *Critique's* first edition, he drew upon the traditional metaphor of metaphysics as "queen of the sciences", but saw her rule as a decadent despotism, ruled capriciously by a factional government lacking legitimacy and therefore constantly under attack. It

is attacked on one side by skeptics, who doubt the authority of any government at all, on another by empiricists, who trace the lineage of the supposed queen to the common rabble of experience and wish to erect in place of the monarchy a republic of science whose foundations, however, Kant regards as inadequate to achieve legitimacy. It suffers a third assault from "indifferentists," who will accept (often uncritically) the non-empirical claims of metaphysics, but refuse allegiance to any "government" with regard to them – that is, they regard such claims as resting on common sense or undisciplined assent, and would deny that they have, or even admit of, any scientific demonstration. Kant's aim, stated in terms of the same metaphor, is to legitimize the monarchy of Queen Metaphysics, while at the same time limiting the scope of its rightful claims to authority (KrV A ix–xii). In the language of this metaphor, the *Critique* is a court of justice, determining the rightful powers of the monarch under reason's natural laws. Thus it is a critique *of* pure reason in both the subjective and the objective sense: it is carried out *by* pure reason *on* the claims made by pure reason. In this way it is also, as Kant says, a Socratic enterprise, aiming – as Socrates did – both at self-knowledge and at the deflation of vain pretensions to know when one does not know.

Or, as Kant redescribes his task in the Preface to the second edition, the aim of the *Critique* is to transform metaphysics from a "random groping" into a genuine science, by limiting it to its proper sphere and grounding it on a rationally well-conceived method (KrV B vii–xiv). Kant's approach to grounding natural science is determined by the conviction that empirical sciences themselves require metaphysical grounding because they must employ certain concepts and principles that cannot be of empirical origin and cannot be justified by appeal to any experience. The concept of causality involves a necessary connection between cause and effect, yet as Hume had shown, no experience of instances in which supposed causes are followed by their effects will suffice to establish more than a contingent conjunction. Mathematical concepts are independent of experience yet their application to the natural world is essential to modern physics; mathematical propositions are clearly *a priori*, not based on experience of their instances, but modern science depends on our knowledge of them.

The general problem of pure reason

Some influential philosophers before Kant (Leibniz and Hume, for instance) had hoped to restrict *a priori* knowledge to propositions expressing "implicit identities" or "relations of ideas" – that is, propositions whose negation involves a contradiction, such as "All bodies are extended" and "Every effect has a cause". Kant calls such propositions

"analytic", because their truth depends on the analysis (or taking apart) of the subject concept. Suppose, for instance, that our concept of body contains three elements – or "marks" (*Merkmale, notae*), as Kant calls them: extension, impenetrability, and shape.[1] Then simply by analyzing this concept, we can know necessarily and *a priori* that every body is extended. Likewise, if our concept of 'effect' contains the marks 'something that happens' and 'produced by a cause', then it is analytic that every effect has a cause (KrV A7–9/B12–13).

Knowledge of analytic truths was regarded on all sides as *a priori* even though the concept of body is empirical, so that experience of bodies is necessary for its acquisition. Kant has no intention of denying that experience is necessary for all knowledge whatever: "All cognition begins with experience" (KrV B1). A proposition is known *a priori* when knowledge of it does not depend in any way on the specific contents of experience, when any experience that would suffice to enable us to entertain the proposition would also be sufficient to give us knowledge of its truth. And this is the case with a proposition such as 'Every body is extended' or 'Every effect has a cause'. For although it may require experience to give us the concepts of 'body' and 'effect', it is not our experience of bodies or effects, but only our own activity of explicating what we think in the concepts of body and effect, that tells us that bodies are extended and effects must be produced by causes.

In Kant's view the epistemic function of analytic propositions is restricted to that of explicating the concepts we use – making clearer or more overt to ourselves what we are thinking in a given concept (KrV A6–7/B10–11). Analytic propositions cannot, therefore, serve as principles in a science, or be used to extend or systematize empirical knowledge, as Leibniz and his followers had thought. *A priori* principles constituting natural science cannot be analytic, because they are not the result merely of discretionary choices about which concepts to use. A principle such as 'Every *change* has a cause', for example, is *synthetic*, connecting the concept of the subject to a predicate lying outside the subject concept, so that the judgment extends or amplifies our cognition of the objects falling under the subject concept. The concept of a *change* is merely the concept of a state of the world succeeded by a different state, and the concept of a *cause* is that of a state of the world upon which a different state follows with necessity, according to a causal law. But it is no part of our concept of a change that the succession of states involved in it is determined necessarily or in accordance with a law. Therefore, if it is to be a part of our conception of the natural world that all changes in it have causes, then our knowledge that this is the case must consist in *a priori* knowledge of a synthetic proposition.

Some mathematical propositions, Kant thinks, are also analytic, such as that 'Every whole is greater than its (proper) part' (KrV B16–17). For

this can be known *a priori* simply by analyzing the concepts of part and whole. But propositions depending on necessary equalities holding in virtue of arithmetical operations such as plus or minus, and geometrical propositions depending on the nature of triangles or circles, are synthetic. Nothing contained merely in the concepts of 'seven', 'five', 'plus', or 'twelve' suffices for us to know that $7 + 5 = 12$ (KrV B15). It is not contained in the concept of a circle (of a curved figure each of whose points is equidistant from a common point) that its circumference is more than three times its diameter, nor is it contained in the concept of a three-sided closed figure that the sum of its angles are necessarily equal to two right angles.

For Kant, *metaphysics* is the science of synthetic *a priori* cognitions through concepts. The traditional problems of metaphysics, those that concern the foundations of the sciences and also those that concern the supposed supernatural questions that concern us, have to do with propositions of which metaphysicians claim synthetic *a priori* knowledge. 'God exists', 'The soul is immortal', and 'The will is free', as well as 'Every change has a cause' and 'Through every change, the quantity of substance in the world remains constant' are propositions of *metaphysics* because cognition of them, if it were to exist at all, could not be empirical, and hence we must aspire to synthetic *a priori* knowledge of them. The "general problem of pure reason" on which the legitimacy of metaphysics rests, is: "How are synthetic *a priori* judgments possible?" (B19).

Kant's answer to this question depends on two theses. The first is an explanation of how *a priori* cognition in general is possible at all. The second is a novel, controversial, and puzzling thesis about the nature of the objects of which we can have synthetic *a priori* cognition which Kant calls 'transcendental idealism'.

How *a priori* cognition depends on our faculties

The *first* thesis is that synthetic *a priori* cognition is possible because what we know of objects *a priori* depends not on them but on our faculties and their exercise. We have cognition of objects at all, in Kant's view, only because they affect us in certain ways, leading to our experience of them (KrV A19/B33). But it does not follow from this that all of what we know of them depends on them and on what experience tells us about them. For in order for us to have cognition of objects, our cognitive capacities must also be engaged (KrV A1, B1). If the operation of our faculties determines the objects of our cognition in certain ways no matter how these objects may be constituted in themselves and no matter what experiential influences they exert on us, then those determinations will belong necessarily to any object we know, or even any object

we can know, and will belong to them *a priori*, that is, independently of any particular experience we might have of them. If the activation of our faculties determines these objects in ways that can be expressed in synthetic judgments about those objects, then the cognitions of objects that depend on our faculties in this way will be *synthetic a priori* cognitions. In that case, the possibility of synthetic *a priori* cognition in general rests on knowledge of what our faculties contribute to experience, and the enquiry that investigates the conditions for the possibility of experience itself will be a new and fundamental branch of philosophy – *transcendental* philosophy.

In the early modern period, there was a dispute over whether some of our ideas are innate – in us at birth, placed there perhaps by God or by our natural endowments as creatures. (Among philosophers still widely studied, "nativism" was defended by Descartes and Leibniz, and famously denied by Locke; Kant was familiar with the doctrine of innate ideas through the writings of Christian August Crusius.) It might be thought that Kant's account of the possibility of synthetic *a priori* cognition is also a form of nativism. Kant seldom mentions this dispute, but when he does (perhaps to our surprise) he flatly denies the existence of innate ideas or innate knowledge of any kind (Ak 2:392–393, 8:221). It illuminates his theory of synthetic *a priori* cognition to see why.[2]

Kant understood nativism as the view that some of our ideas or cognition is supplied to us from a source other than our senses (for example, from God or nature). He rejects this as an account of how *a priori* cognition arises because it treats some of our knowledge as consisting in a species of non-sensory *data*, as something given to us in some way other than experience. Innate ideas, as Kant understands them, would be thoughts that come not from our thinking but lie in us "pre-formed" – as (according to some then current theories of the reproduction of living things, which Kant also rejected) a miniature version of each organism is present "pre-formed" within the sperm of its male parent (KrV B167; cf. KU 5:421–424). Such a theory, he thinks, would attribute the very form of our thoughts not to us but to some pre-determining (divine or natural) power. Instead of this, Kant regards everything *given* to us as given *empirically*, through sensation: our knowledge, which falls entirely within the bounds of experience, is nothing but what we make of these data through our passive capacity to receive it and our active ability to organize it. In other words, Kant thinks of *a priori* cognition as nothing but *our* exercise of our faculties on what is given empirically, *our* active contribution to the knowing process. A *Critique of Pure Reason*, as a project of self-knowledge, is possible because we ourselves, as rational beings, are the creators of "pure reason" and are therefore capable of understanding both what we do and how we do it.

What must the objects of *a priori* cognition be like?

If we are to regard the objects of our cognition as determined in some respects by our active exercise of our cognitive faculties, then how must we think of these objects in order to understand their properties as determined in this way? The answer to this question leads us to Kant's *second* crucial thesis about synthetic *a priori* cognition, which is his famous (or notorious) doctrine of *transcendental* (or "critical") *idealism*. This doctrine says that we have cognition only of "appearances," not of "things in themselves". The objects of experience are empirically real, but transcendentally ideal. This new way of thinking about the objects of our empirical cognition is, according to Kant, necessary if we are to answer the question: "How is synthetic *a priori* cognition possible?" And Kant compares the revolution in thinking necessary to embrace this doctrine to the revolution in thinking required to embrace Copernicus' heliocentric theory of celestial motion (KrV B xvi). Before Copernicus, we thought that heavenly bodies moved but we earthly observers were at rest. Now we see that we observers too must be regarded as in motion. Analogously, before Kant we thought that our cognition depended on its objects; but now we must see that the objects we cognize must depend on our mode of cognizing of those objects. In both cases, we made an assumption that was natural because our attention was focused on the objects of our knowledge and not on our own relation to them. Hence everything seemed to depend on the objects we observed and not on us. The revolution in both cases consisted in taking account, contrary to the way things naturally seem, of our own role in the processes we are trying to observe and understand.

Ever since the publication of the first edition of the *Critique of Pure Reason*, this doctrine has been a source of perplexity and controversy. Some of Kant's earliest readers saw it (to Kant's chagrin and even in the face of his strenuous denials) as a species of Berkeley's idealism, a metaphysical position that denies the reality of material objects and even of space (perhaps even of time). Many others have regarded Kant's denial that we can know objects "in themselves" as a capitulation to extreme skepticism. Still others thought that Kant's position can be made defensible (or even internally consistent) only if the existence of "things in themselves" (perhaps even the very intelligibility of the notion) is utterly denied. In chapter 4, we will see that there are some good reasons, lying in Kant's very own statements of transcendental idealism, for puzzlement about what this revolutionary doctrine actually maintains. For the present, however, it will be best to set these questions aside and look instead at Kant's more detailed account of how our knowledge arises – and of the way the operation of our faculties makes synthetic *a priori* cognition possible.

Intuitions and concepts

Early modern epistemology recognized two main sources of knowledge: sensation and thought. But philosophers differed regarding the role of each in knowledge and especially regarding how fundamental each is as a source of knowledge. Descartes recommended that we distrust the senses and rely on the proper use of our intellect in acquiring knowledge. Even more importantly, he regarded sensation itself as a species of thinking – a species whose lack of clarity and distinctness makes it inferior as a source of knowledge. Leibniz formulated a similar idea succinctly when he declared sensation to be "confused" thinking: thinking whose content is inherently imprecise. Spinoza similarly regarded sensations as an inferior grade of knowledge, in which the mind is relatively opaque to its own activities and operations. Locke, on the other hand, treated all our ideas as having their source either in the senses or in our reflection on the operations of our minds in dealing with what the senses give us. More radically, Hobbes treated thoughts as nothing but the faint remains left in us by sense impressions. Hume classified thoughts (or "ideas") as fainter copies of 'impressions', of which sensations constitute the most common and obvious class.

Kant rejects the attempt either to classify sensation as a species of thought or to explain thought from sensation. Instead, he sees sensations and thoughts as performing distinct cognitive functions, and regards genuine cognition as occurring only through their thoroughgoing combination, resulting from the co-operation of the faculties pertaining to each of them. Cognition of an object requires, namely, that the object be somehow (directly or indirectly, actually or virtually) *given to* the mind, and also that the mind actively combine our representations in such a way as to make possible a genuine *experience* of the object, grounding *judgments* about it and also the combination of such judgments in inferences and structured theories displaying the coherence of that experience. The receptivity of the mind that enables an individual object to be given to cognition Kant calls *intuition*; the active function of mind enabling representations to be combined he calls *thinking* (or *conception*).

In beings like us, intuition occurs through the effect an object has on us, and our capacity to receive such effects is called *sensibility* (KrV A19/B33). Kant's use of the term 'intuition' has nothing to do with any quasi-mystical connotations this word may have in ordinary English. The German word *Anschauung* simply means 'looking at', and the Latin word *intuitus* (which Kant regarded as its equivalent) was the traditional term used in scholastic epistemology for any immediate cognitive contact with individual objects. 'Intuition' (in Kant's usage) is ambiguous: it can refer to the state of being in such contact, or to the thing with which

we are in contact regarded simply as an object of intuition, or to the mental state (or representation) afforded us when we intuit an object. For Kant, all our intuition comes through the senses, which put us in immediate contact with objects through the influence they have on us. Objects are thus given to us in intuitions through sensations. An intuition is a species of representation resulting from an object affecting us that puts us in an immediate cognitive relation to the object.

Representations (whether they are sensations or any other sort of mental item) afford us genuine cognition, however, only if they are combined in such a way as to stand for the objects that cause them and enable us to make judgments predicating properties of these objects. When I sense an apple's red color, smooth texture, or tart taste, I obtain cognition of the apple only if I can form the judgment that properties falling under the concepts 'red', 'smooth', and 'tart' are truly predicated of the apple as a thing falling under the concepts 'material object' or 'fruit'. My sensations, however, do not directly provide me with the concepts in terms of which to form such judgments. In fact, there are many different ways in which subjects are capable of conceptualizing what is given to them in sensation. Conceptualizing and judging depend on how the subject combines its representations or relates them to one another – for example, whether a subject regards the apple as a single piece of fruit or instead a collection – for instance, of living cells or organic molecules.

This combination of representations is not an act of intuiting, nor is combination any work of our (passive) sensibility at all. All combining is done by our *active* cognitive faculties. These include *understanding* (which forms concepts of objects and makes judgments about them) and *reason* (which connects such judgments through inferences and unifies our cognition under *principles* specifying the self-directing aims of our cognitive faculties as a whole). Neither understanding nor sensibility is reducible to the other. Although understanding is "higher" than sensibility in the sense that it has the authority to order what is given in sensibility, neither faculty is to be "preferred" to the other as a source of knowledge, since for any cognition both are required and both must work together. "These two faculties or capacities cannot exchange their functions. The understanding cannot intuit anything and the senses cannot think anything." Nor can either produce cognition without the co-operation of the other: "Thoughts without content are empty, intuitions without concepts are blind" (KrV A51/B75).

In an abstractly conceived divine understanding, objects are intuited through the active faculty (the understanding) that also creates the objects known; and in the absolute simplicity of God's nature and action there is no need (or even any possibility) of combining representations, or inferring one judgment from another. God is impassible (hence he has no senses) and his intellect is active by creating, not combining, so he

has no need for thoughts or concepts to organize what he knows. In finite creatures like us, however, who cognize objects through being affected by them and then organizing this data in thought, two heterogeneous faculties of cognition are needed.

Kant's dualism about our cognitive faculties was an object of suspicion and skepticism on the part of his earliest readers and followers. Some thought that his apparently unargued assumption that we are affected by objects outside us begs the question against important species of skepticism and idealism to which Kant should have furnished a more substantive reply. Others wondered how Kant, who proposes critically to examine our faculties of knowledge in order to estimate in advance what we can and cannot know, can claim any knowledge at all about these faculties themselves. They insisted that a critique of reason, or transcendental philosophy itself, requires a "meta-critical" investigation of how we know these faculties, which should put in question Kant's own "dogmatic" claims about them. Of course these very problems, and the creative theories generated in response to them by such philosophers as Reinhold, Fichte, and Hegel, were unthinkable before Kant's project itself made it possible to articulate them. It cannot be said that Kant himself addressed them directly anywhere in his writings. Defenders of Kant should not claim that his immediate followers were on the wrong track in asking for solutions to them.[3] But it may help make Kant's starting point more intelligible to attempt to describe his project in the *Critique of Pure Reason* in such a way as to make it clear why he might not have thought he had to worry as much about them as did some of his critics and immediate successors.

Perspectivity and objectivity

We all take our experience to be perspectival, in the sense that each of us knows the world from a single point of view, which is to be contrasted with the point of view of other actual or possible experiencers. But we also make judgments about the objects we experience that we think of as sometimes actually (or at least possibly) *true* – implying that no matter what perspective another experiencer might occupy, it can judge correctly only if it agrees with those judgments. The fact that I experience only from my unique individual perspective derives from the fact that my direct contact with individual items of experience is peculiar to me, whereas my capacity to formulate judgments claiming truth (for everyone, no matter what their perspective on the world) depends on the fact I can think about what I experience in a way that is not wholly tied to my perspective.

This same contrast can be seen as a peculiarity about the concept of the 'I', of that which is the subject of experiences. 'I' is always used to

refer to the occupant of one particular perspective, in contrast to other possible occupants of other possible perspectives. If we did not grant that other perspectives besides mine are at least possible, then there could be no meaningful reference to me as the subject of this experience or to this perspective as "mine". At the same time, the concept 'I' has the contrasting peculiarity that it is capable of referring to the subject of any perspective at all – any experiencer whatever (or even any possible experiencer) counts as an 'I', since only the fact that it is an 'I' makes it the occupant of a possible perspective at all. Thus 'I' is both the most singular and the most universal term, and it is inevitable that the fact that it is the one makes it also the other. For only if the same concept can simultaneously play the role of the occupant of this perspective and the role of the occupant of any possible perspective is it possible for there to be a plurality of possible perspectives on one and the same reality about which (possibly) true judgments can be made.

This duality of 'I', corresponding to the duality of the perspectivity that belongs to our experience of a single reality, is the ground of Kant's duality of intuition and thinking. In Kantian parlance, it is *intuition* that represents the immediate, individual contact between knower and object that makes perspectivity possible, while *thinking* is what makes possible the concepts that afford to the occupant of any possible perspective the opportunity of making judgments that are true, and hence valid equally for all perspectives.

Perhaps it is possible for a certain kind of radical skeptic to question even the possibility of a plurality of perspectives, and especially the reality of a world of objects about which non-perspectivally true judgments can be made. Yet it is difficult for a skeptic to question this self-consistently, since any judgment that it is false, or even doubtful – any denial or doubt that there are judgments that are (non-perspectivally) true, hence valid for all perspectives – necessarily makes an affirmative claim to the very same truth that the skeptic is attempting to call into question. So the only sort of skeptic who could question this is one who, following some of Sextus Empiricus' suggestions, truly abstains from all assertions or judgments whatever, even from the judgment that anything is doubtful. Taken strictly, in fact, it would be impossible for such a philosopher, who makes no assertions or judgments at all, even to dispute or doubt anything, since doubting is a doxastic stance that makes sense only when it stands in contrast to the possible assertion that the doubted proposition is true or certain, and therefore amounts to a commitment to *assert* and *judge* that the proposition is doubtful rather than certain.

Even if such skepticism were conceivable, Kant would not be without resources in combating it. Kant does begin the transcendental analytic with an account of cognition as requiring intuition and thought, the

givenness of objects and the connection of representations into a whole of cognition. But this account is *not* assumed dogmatically and simply appealed to in the course of his subsequent arguments (as some of his early skeptical critics like to pretend). The method of transcendental philosophy is very far from permitting any such thing. In the Transcendental Aesthetic, Kant tries to provide an account of our very concepts of space and time that reveals them as necessary forms of intuition. The perspectivity of our experience has its root in the fact that every experiencer is located at one position in space, that every experience is determinately oriented from that position, and takes place at a determinate now in time, with an orientation to other already actualized times determinately ordered as past, and other still non-actualized times are related to as a similarly determinate order of possible futures. If, as Kant argues, the very concepts of space and time reveal space and time to be forms of our intuition, then someone who tries to doubt that we have intuitions of space and time would also have to reject concepts presupposed by all thinking that regards itself as starting from a here and a now and directing itself toward a there and a then. That would make it impossible even to *ask*, for instance, whether there is a world external to our experiences or whether the future will be like the past. It would thus render skeptical doubt about these matters unintelligible.

In the Transcendental Deduction, Kant attempts to argue that any subject that deems itself capable even of grasping a temporal series of its own subjective representations – *as* a series, for *this* subject – must be able to conceive those representations in such a way that they refer to items that count as "objects," about which it can make judgments claiming validity for any possible subject of experiences. In the Principles of Pure Understanding, Kant further presents a connected argument showing that for there even to be an experienced series of subjective experiences, these experiences must have objects that are regarded as constituting a world of substances governed by causal laws distinct from any of my subjective representations. Again, a skeptic who will not grant even that we can apprehend a series of our subjective experiences through time has no way even of self-consistently *formulating questions* about whether we can cognize objects distinct from these experiences or whether apparent changes in such objects are real changes or are governed by causal regularities. The effect of Kant's arguments is therefore to show that skeptics can doubt the existence of an ordered empirical world of causally interacting substances distinct from the subject's own mental representations only on pain of undermining the intelligibility of what must be presupposed even to make sense of their own questioning.

Such anti-skeptical arguments in the Transcendental Aesthetic and Transcendental Analytic do not depend on assuming the truth of Kant's

initial picture of the duality of intuition and thinking, or of experience as resulting from the effects of real objects on our sensibility, or on the assumption that we can have knowledge of these objects that is valid for every conceivable perspective. Instead, they begin with assumptions about our experience that are so thin and undemanding that a skepticism which questioned them would undermine its own philosophical interest, perhaps even its own intelligibility. Kant's strategy is to show that what the skeptic wants to doubt is presupposed even by a concept of experience that must be assumed to make sense of these doubts themselves.

II the transcendental aesthetic: pure sensibility

The truth perceived by those who raised 'meta-critical' doubts about Kant's project is rather this: that along with his anti-skeptical arguments, Kant is constantly attempting to integrate their results into his theory of experience as objective, yet perspectival, and therefore as resulting inevitably from the operation of our understanding on data provided by sensible intuition. The Transcendental Aesthetic and Transcendental Analytic are therefore not *only* (perhaps not even chiefly) exercises in anti-skeptical argument, but they *also* contain a positive theory concerning the necessary constitution of our experience and the operation of our cognitive faculties.

Space and time

Kant's discussion of space and time in the Transcendental Aesthetic takes place against the background of the controversy between the Leibnizian and Newtonian views, as they were discussed in the correspondence of 1715–16 between Leibniz and Samuel Clarke (the British arch-rationalist whom Newton chose to represent his position in the controversy with Leibniz). The Newtonian position is that space and time are actual entities, existing independently of our minds and of the material objects that occupy them; Leibniz's view is that space and time are conceptual constructs made by our minds as a way of systematizing the perceived relations between things (relations such as 'three seconds earlier than' and 'two meters to the left of'). Kant found both accounts unsatisfactory because they cannot account for the apriority of geometry as a cognition of space, and the *a priori* knowledge of both spatial and temporal quantities that we have in arithmetic. For both accounts treat our knowledge of space and time as dependent on our acquaintance with independently existing things or their properties, and this acquaintance can be only empirical, never *a priori*.

Kant also thought that both accounts violate our very concepts of space and time in crucial ways. It is contained in our concepts of space and time that space and time are singular entities, immediate acquaintance with which is presupposed by our acquaintance with any of the things that occupy them, or even of the spatio-temporal properties of such things (KrV A23–25/B38–41, A30–31/B46–49). On the Newtonian view, space and time are real things, but they have no directly observable properties, so they are not the kind of real things with which we could be immediately acquainted (as we take ourselves to be immediately acquainted with space and time in every experience). The Newtonian thus leaves it hopelessly mysterious how we could ever gain the kind of cognitive contact with space and time that belong to the very concepts we have of them.

Likewise, the perception of the relations Leibniz proposes to use in conceptually constructing space and time is possible only on the basis of prior acquaintance with space and time *themselves*. For it is only by being directly aware of the passing of time that we could ever perceive that one happening occurs three seconds before another, and it is only by perceiving objects as already in the one comprehensive space that we could even understand what it means for one object to be two meters to the left of another. The problem with both theories, in Kant's view, is that they try to grasp space and time as if they were like the real things, properties and happenings that we experience as occupying them. But our very concepts of space and time, he thought, preclude their being anything like these.

Kant's radical proposal about the nature of space and time is that they are *forms of intuition*, necessary ways in which cognizers like ourselves can make cognitive contact with things. Neither they nor the spatio-temporal properties of things and events have any existence apart from our capacity to intuit objects and changes in them. This proposal can be seen as arising from the insight that the awareness of temporality (of being located now, at this determinate moment of time) is fundamental to the perspectivity of every experience that is possible for us, and that being positioned and oriented in space is equally fundamental to the ineluctable perspectivity of our experience of anything we take to be other than ourselves and our subjective experiences. Space and time are therefore not among the independently existing objects given in experience, nor are spatio-temporal properties of things *fundamentally* properties of such objects. Instead, space and time fundamentally have to do with ways *we relate to objects* when we intuit them – that is, when we come into immediate cognitive contact with them on the basis of our unique cognitive perspective. The spatio-temporality of the world is not fundamentally an objective feature of it, but a function of the perspectivity of our experience of it. What is objective about space and

time depends on what is common among all subjects regarding that perspectivity.

Space and time, however, are also *formal intuitions*, that is, they are like intuited *objects* in that they are individuals with which our intuition places us in an immediate cognitive relation. Space and time are like objects in that any intuition of an object is also an intuition of the space in which the object is located and of the time in which the experience itself takes place. Space and time, Kant insists, are also intuited as *single* objects, in that any awareness of one location in space (say, of this room) is also an awareness of *space as a whole*, to which the space in this room belongs (as a proper part); and any awareness of an instant in time (the present instant) is an awareness of the whole of time in whose flow this instant occurs. Space and time are formal intuitions in that they condition our intuition of all the objects we intuit by grounding a system of relations pertaining to objects and the changes that can happen to them, within which those objects have the properties that can be predicated of them, and within which those properties can alter. Spatio-temporal locations of things and happenings within this system obtain not merely within a given subject's perspective but objectively, from every spatio-temporal perspective – just as if space and time themselves were independent objects – even though they are not. Space and time do not merely *consist* (as Leibniz thought) in such an intellectually represented (conceptually constructed) system of relations, since our original relation to them is not intellectual (conceptual) but intuitive. But Kant does follow Leibniz in thinking that the objective placement of an object (in space) or an event (in time) is determinable only relative to other objects and events – there are no 'absolute' positions in space or time, as the Newtonians thought.

Pure sensibility, *a priori* intuition

Because space and time pertain not to objects but to our faculties of intuition, and because they are necessary conditions of any intuition, our intuition of them is *a priori* – that is, this intuition is independent of the particular content of the sensations we receive from any of the real things we intuit in space and time. When the relations of things within space and time are conceptualized, Kant holds that they make possible a system of *a priori* knowledge that applies necessarily and universally to all objects that can be intuited. This knowledge constitutes pure mathematics – geometry as an *a priori* science of any possible physical space, and arithmetic as the science of quantity, applying *a priori* to any magnitude (whether spatial or temporal) found in the things that occur in space and time. Kant's theory that space and time are pure forms of intuition belonging to our cognitive capacities rather than independently existing

entities (or conceptual constructs based on the properties of such entities) seemed to him the only way of explaining the possibility of the synthetic *a priori* cognition we find in mathematics.

Of course there have been major changes in geometry and in physics since Kant's time. For him, the only geometry was Euclid's, and it was a given both that Euclidean geometry gives us *a priori* knowledge and that this is knowledge directly about *physical* space. But the nineteenth and twentieth centuries saw the development of non-Euclidean geometries, and it became an empirical question which geometry describes physical space. So Kant's treatment of these matters, though ingenious and cogent in his time, is no longer applicable in ours, as many have observed. But since Kant's time there has also occurred another change, even more radical, that is less often remarked upon. For a philosopher in Kant's day, it still seemed possible to formulate a single theory of space and time that both accounts for our knowledge of them in mathematics and physics and also accounts for the way space and time, as part of our lived experience, are fundamental to our everyday awareness of the natural world as well as our scientific knowledge of it. Today the mathematical models offered by physicists may account for the data and a whole range of bewildering theoretical constraints required to systematize it, but they cannot be said to provide any interpretation at all of space and time as we live them in sense-experience and as they play a fundamental role in our everyday awareness of the world. Kant's theory, therefore, though no longer tenable, is of continuing interest because it is one of the last plausible attempts to achieve a unified theory of space and time simultaneously from both a scientific and an everyday epistemological perspective. This is something we still need, even if we do not know any longer how to obtain it. Physics itself will never again be an intellectually satisfying department of knowledge until we somehow regain a conception of space and time that reconciles science with everyday lived experience.

The transcendental ideality of appearances

The most radical conclusion from Kant's theory of space and time is one he hastens to emphasize. If space and time are neither things existing apart from our intuition nor relations between the properties of such things, then they are, as Kant says, *appearances*, having no existence in themselves. But what then of the *objects* that appear in space and time? They too are appearances, which we cognize not as they may be in themselves but only as they can be intuited by us. Kant always regarded his most basic argument for the transcendental ideality of the objects we can know as based on the transcendental ideality of the space and time in which alone such objects can be cognized by us. Kant insists,

however, that the ideality of space and time, and the objects in them, is only *transcendental* – that is, it refers to the status they have within a theory telling us how our experience is possible. Empirically, space and time, and spatio-temporal objects in them, are real. They are not illusions; they are to be distinguished from what, empirically speaking, we call 'mere appearances' (dreams, hallucinations, mirages, and the like). Their empirical reality, however, does not consist in their being things existing in themselves, independently of the conditions under which we cognize them. It consists rather in the way they conform to an order of nature that, Kant will argue, is transcendentally necessary if experience is to be possible at all. Whether appearances even *have* an existence in themselves is a question to which Kant unfortunately gives us no self-consistent answer. I postpone until chapter 4 the discussion of the perplexities of transcendental idealism arising from this ambivalence.

III the transcendental analytic: pure understanding

Transcendental logic

Objects are *given* to our cognition through sensible intuition, yielding sensations, and whatever is cognized through sensation must be in space and time, and our cognitions of space and time yield a body of synthetic *a priori* cognition through intuition – *pure mathematics*. But no cognition, not even anything worthy of the name 'experience', would be possible merely in the form of unconceptualized sensations. "Intuitions without concepts are blind" (KrV A51/B75). Cognition requires also the activity of thinking, conceptualizing, and judging, which is performed by our faculty of understanding. Thus in addition to the Transcendental Aesthetic, which deals with the *a priori* conditions of experience as forms of intuition (space and time), a critique of pure reason must also contain a Transcendental Logic, which deals with the *a priori* conditions of thinking.

"Logic" is for Kant a science dealing with the use of the understanding in thought. Traditional formal logic (first devised by Aristotle and then codified by the scholastics) abstracts from all content of thinking, and especially from the way thinking relates to objects of cognition. But Kant argues that the objects of cognition, as appearances, are constituted not only by the way we intuit them in space and time but also by the way the understanding thinks of them as *a priori* conditions for the possibility of experience in general. The science which deals with objects in this way is not formal (or general) logic but transcendental logic.

One part of Transcendental Logic, which Kant calls the Transcendental Analytic, deals with conditions of the possibility of experience that yield synthetic *a priori* cognitions through concepts – *metaphysics*. The Transcendental Analytic in effect aims at expounding and justifying *a priori* a conception of the empirical world. This is a world of material objects in space that change over time, whose properties and changes are subject to mathematical measurement. These objects are all constituted by a substance whose determinations alter but whose quantity does not increase or decrease. Different material objects are distinguished from one another by the placement of distinguishable parts of this substance in different regions of space. Their relationships to one another at any given time, and also their alterations through time, are governed by necessary causal laws. Material substances are distinct from our inner or subjective representations – they constitute a real or external world. The judgments we make about them, their states and their relations, seek, and can in principle lay claim to, objective validity, validity for "consciousness in general" – that is, for all beings capable of making such judgments at all. The project of the Transcendental Analytic is therefore to vindicate *a priori* the fundamental conception of the world to which modern mathematical natural science is committed, and to respond to skeptical objections to the reality and knowability of that world.

The other part of Transcendental Logic, called Transcendental Dialectic, deals with principles deriving not from understanding but from reason. Kant thinks the latter principles are indispensable guides for empirical inquiry and for systematizing the results of that inquiry, but he also thinks that reason tends, in its own operation, to be led astray by its principles – it is subject to a kind of illusion that makes it think the principles and self-made concepts that should direct its empirical inquiry also afford it metaphysical (or synthetic *a priori*) cognition of non-empirical objects (such as God, free will, and an immortal soul) that cannot be given to beings like ourselves in any intuition. 'Dialectic' for Kant means a "logic of illusion"; it is the purpose of the Transcendental Dialectic to expose the dialectical illusions in metaphysics, correct its errors, and assign the legitimate rational principles leading to these illusions to their proper sphere and function. In chapter 5, we will discuss the Transcendental Dialectic. The remainder of this chapter will be devoted to an (all too brief) exposition of the famous and complex chain of argument found in the Transcendental Analytic. The aim of this argument is to show, contrary to the skeptics, that the constitution of the objective world, the fundamental order of nature as modern science conceives and investigates it, is necessary *a priori* as a condition for our having any experience at all.

This famous and highly ambitious argument is divided into three main phases. The first phase, which Kant (once, at KrV B159) calls the

'metaphysical deduction [of the categories]', proposes to identify twelve fundamental concepts, called 'categories', whose origin is *a priori* in that we acquire them not from the experiences to which we apply them but from the exercise of our own understanding. The second phase, called the 'transcendental deduction of the categories', aims at showing that despite their origin in our faculties rather than in objects of experience, the categories necessarily apply to any objects that can be given to our senses. The third phase, contained in a chapter entitled "System of all the Principles of Pure Understanding," argues that the categories necessarily apply to experience in determinate ways: for instance, that both the spatio-temporal forms and the real qualitative contents of experience are quantitatively measurable and that all changes we can experience are alterations in the determinations of an underlying substance and that they follow causal laws.

As I have just said, the arguments of the Transcendental Analytic are best viewed as a single lengthy but interconnected chain of argument, the earlier links of which are in some measure dependent on the later ones, for both their intelligibility and their persuasiveness. If there is a phase of the argument that is fundamental, then it is probably (as Kant himself suggests, KrV A xvi) the transcendental deduction of the categories. For the metaphysical deduction uses the logical form of judgments as its guideline for the discovery of those concepts that are both *a priori* and fundamental, and it is only in the transcendental deduction that Kant argues that any possible experience must contain objects about which any subject of experience must form judgments laying claim to universal validity for consciousness in general (for all subjects of experience). And it is also the applicability of the categories to all objects of the senses (established in the transcendental deduction) that provides a guarantee that what is given to the senses must be determinable through the categories of quantity and quality (grounding the mathematical principles of understanding, KrV A162–176/B201–218), as well as that time, as the form of all intuition, is "determined" in respect of duration, succession, and co-existence, which grounds the analogies of experience (which we will explore further in the next chapter), and the laws of substance, causality, and reciprocity (KrV A176/B218).

The metaphysical deduction: forms of judgment and categories

Kant's point of departure is the conception of a judgment in traditional (scholastic-Aristotelian) logic. In this tradition, the general form of judgment was:

S is P

that is, *a subject term* connected to a *predicate term* by means of the *copula* 'is'. Different possible kinds of judgment can therefore be displayed by analyzing the kinds of subject term, the kinds of predicate term, and the kinds of copula. For example, all subject terms are of three kinds:

Universal: *All* S
Particular: *Some* S
Singular: *The* S

Predicates form a similar triad:

Positive: P (a positive property)
Negative: not-P (the negation of a positive property)
Infinite: non-P

An infinite judgment is one in which the predicate is a positive property, but one that is signified by a concept that restricts part of a domain of such properties by negating the remainder of the domain. For example, I could designate the color of the book before me by saying that the book is non-red (meaning that it is either yellow, or blue, or some other color, but not red). This differs from the negative judgment because a negative judgment carries with it no implication that the predicate is any positive property at all. For example, it would be entirely correct to predicate of the number three that it is not-blue, but incorrect to apply to it the infinite judgment "The number three is non-blue," because this would be to assert (what is false) that it has some other color.

There are similarly three kinds of copula, corresponding to the modal status of a judgment:

Problematic: S *is possibly* P
Assertoric: S *is (actually)* P
Apodictic: S *is necessarily* P

In addition, there is a fourth triad arising from the fact that in traditional logic judgments may be combined with one another in syllogistic inferences, and this in three ways:

Categorical: All S is P (and All P is R; therefore, All S is R).
Hypothetical: If S is P, then S is R (and S is P; therefore, S is R).
Disjunctive: S is either P or R (and S is not R; therefore, S is P).

This gives rise to a "Table of Judgments" (KrV A70/B95) that Kant arranges as follows:

Quantity of Judgments
Universal
Particular
Singular

Quality
Positive
Negative
Infinite

Relation
Categorical
Hypothetical
Disjunctive

Modality
Problematic
Assertoric
Apodictic

Kant's next claim is that corresponding to this table is a "Table of Categories" (KrV A80/B106), each corresponding to an entry in the table of judgments. These twelve *a priori* concepts are such that their instantiation in the sensible world makes experience possible:

1
Of Quantity
Unity
Plurality
Totality

2
Of Quality
Reality
Negation
Limitation

3
Of Relation
Of inherence and subsistence
(*substantia et accidens*)
Of causality and dependence
(cause and effect)
Of community
(reciprocity of agent and patient)

4
Of Modality
Possibility–Impossibility
Existence–Non-existence
Necessity–Contingency

Some of these correspondences are self-evident, but others involve subtleties for which we have no space here.[4] For our purposes, the most important point in the general relation between each category and its corresponding judgment form is that our faculty to judge according to that form carries with it the capacity to organize our representations under the corresponding concept. But Kant is not committed to (and does not) hold that whenever we employ a judgment form, we thereby employ the corresponding concept. (For instance, when I say 'If this book is red, then it is colored', I do not have to think that the book's being red is the *cause* of its being colored.) His claim is rather that what enables us

to form concepts like 'one', 'many', 'all', 'substance', 'cause', and so forth is simply that we are able to judge the things we experience in accordance with these judgment forms. The categories are not like empirical concepts, such as 'red' or 'dog' or 'can-opener' that might either apply or fail to apply to our experience, depending on what its sensory contents might be. Instead, we bring these concepts to our experience (along with the forms in terms of which we make judgments about it). Empirical concepts themselves are always instances of categories (a dog is 'one' animal, and a substance; 'red' is an accident of a substance and a *reality* (or positive property); a can-opener is a substance and also a thing with the causal capacity to open cans of soup). In any array of data about which we can make judgments there will necessarily be instances of 'one', 'many', 'all', and so forth.

One of the earliest criticisms of Kant to gain currency – this was one important part of the "meta-critical" problem mentioned above – was that he was taking too much for granted in assuming that the forms of judgment in traditional logic may be taken, without further justification, to represent the only way we might conceptualize our experience. The history of Kantianism has seen many versions of this worry, and many attempts to turn Kant's forms of intuition and table of categories into a conceptual structure that might be historically alterable and dependent on changes in the history of science or on the choices of investigators about how to conceptualize their inquiries or their subject-matter. The obvious worry here, of course, is that allowing such modifications of Kant's project would jeopardize either the syntheticity or the apriority of the principles whose synthetic *a priori* status is required if Kant's limited defense of metaphysics is to be preserved against skeptical objections. Such controversies continue down to the present day, and the way they have reshaped both philosophy and scientific inquiry are an important part of Kant's enduring legacy.

further reading

Henry Allison, *Kant's Transcendental Idealism*. New Haven: Yale University Press, 1983.

Michael Friedman, *Kant and the Exact Sciences*. Cambridge, MA: Harvard University Press, 1992.

Sebastian Gardner, *Routledge Guidebook to Kant and the Critique of Pure Reason*. London: Routledge, 1999.

Paul Guyer, *Kant and the Claims of Knowledge*. New York: Cambridge University Press, 1987.

Patricia Kitcher (ed.), *Kant's Critique of Pure Reason: Critical Essays*. Lanham, MD: Rowman and Littlefield, 1998.

Béatrice Longuenesse, *Kant and the Capacity to Judge*. Princeton: Princeton University Press, 1998.

Wayne Waxman, *Kant's Model of the Mind*. New York: Oxford University Press, 1991.

notes

1 It is sometimes thought that analytic propositions must be 'true by definition'. But this is not a formulation that Kant would countenance. He thinks that a genuine definition of a concept must involve a demonstration that the concept can apply to something, and it must exhibit the definiens as presenting the complete content of the concept. Definitions meeting these requirements can, in his view, be found only in mathematics, where we can construct our concepts *a priori* by exhibiting their objects in pure intuition (KrV A727–732/B755–760). Empirical concepts, such as 'body' or 'bachelor', have no definitions, because there can be no demonstration of their possibility except the contingent actuality of the experiences from which we draw them, and because their limits are inherently vague: further experience may always lead us to think more in them than we did at the start. These considerations, which are closely akin to those leading W. V. O. Quine to put the entire analytic/synthetic distinction into question in the mid-twentieth century, did not seem to Kant to show that we cannot be certain of some of the marks belonging to a given concept, and so they did not seem to him to show that there could not be analytic judgments made on this basis. See Quine, "Two Dogmas of Empiricism," *From a Logical Point of View*, 2nd edition (Cambridge, MA: Harvard University Press, 1980) and Philip Kitcher, "How Kant Almost Wrote 'Two Dogmas of Empiricism' (And Why He Didn't)," *Essays on Kant's Critique of Pure Reason*, eds. J. N. Mohanty and Robert W. Shahan (Norman: University of Oklahoma Press, 1982), pp. 217–49.

2 For Kant's rejection of nativism, see Graciela de Pierris, "Kant and Innatism," *Pacific Philosophical Quarterly* 68 (1987): 285–305, and Lorne Falkenstein, "Was Kant a Nativist?" in Patricia Kitcher (ed.), *Kant's Critique of Pure Reason: Critical Essays* (Lanham, MD: Rowman and Littlefield, 1998), pp. 21–44.

3 As is argued by Karl Ameriks in *Kant and the Fate of Autonomy: Problems in the Appropriation of the Critical Philosophy* (New York: Cambridge University Press, 2000).

4 The best discussion of these matters is Béatrice Longuenesse, *Kant and the Capacity to Judge* (Princeton: Princeton University Press, 1998).

chapter 3

the principles of possible experience

I the transcendental deduction of the categories

Skeptical doubts successfully challenge the possibility of our knowledge only as long as the skeptic implicitly grants us certain things that we naturally think constitute a sufficient foundation for what we think we know, and then surprises us by showing that those things give us less adequate justifications for our supposed knowledge than we thought they did. This is how skeptical doubt operates, and also indicates the limiting conditions under which skeptical arguments can retain their philosophical interest.

For example: We are aware of having a series of mental states through time that exhibit certain consistencies, recurring similarities, and constant conjunctions. On the basis of these patterns, we judge them to be caused by a world of objects outside us, by enduring substances whose perceptible changes follow certain causal regularities. But now skeptical worries intervene. We have no access to the causal connections between our states and whatever it is that may produce them, so we really have no basis for our causal judgments about their origin. Nor are constancies and recurring patterns firm evidence for either the identity of substances through time or for the necessity of a sequence between events that is required to justify a claim that there are causal connections involved. Where we thought our experiences gave us the evidential equivalent of cold cash, these skeptical doubts leave us feeling like we are holding instead only a handful of yesterday's losing lottery tickets. The skeptical problem is not constituted by the fact that there might be substances or causal connections other than the ones we think there are. For our experience does give us evidence on this point, and we don't claim infallibility about such judgments anyway. The real problem is that the skeptical argument calls into question whether we could be warranted in applying concepts such as 'object', 'substance', and 'cause' to what is presented us no matter what the evidence of our experience might be like.

Kant compares the task of the philosopher here to the task you would face as a prosecutor (in a Roman-based legal system) who is trying to make a legal case against a defendant (KrV A84/B116). You have two distinct questions to settle:

1 The *quaestio quid iuris* (the question of right – that is, of what right under the law you claim the defendant has violated – which is equivalent to showing that the charge against the defendant has a valid legal basis): For this you must derive from the legal statutes a proposition of the form: "If the defendant did X, then he is guilty of crime Y." (For instance, if he *removed property that was not his*, then he is guilty of *theft*.)
2 The *quaestio quid facti* (or question of fact): You must present evidence that the defendant did X (that he removed property that was not his).

Both points must be proven if the case against the defendant is to be made. For if the defendant can't be shown to have done X, then he is innocent of any crime whatever that supposedly involved his doing X. And no matter what the facts are, even if he did do X, he still committed no crime if it cannot be derived from the statutes that doing X constitutes committing some crime Y.

Kant claims already to have shown that concepts such as *substance* and *cause* are *a priori*. They come not from experience but from the forms of judgment employed by the understanding. With respect to *a priori* concepts such as *substance* and *cause*, however, there is a possible "question of fact" which concerns the experiences through which we first encountered instances of these concepts or that might entitle us to employ them in particular cases. The question of right, however, concerns our entitlement to use these concepts at all, no matter what the facts might be. The technical legal term in Roman law for argument leading to the answer to the question of right is the *deduction*. It is in a metaphorical sense based on this usage that Kant coins the term 'transcendental deduction'. Because Kant uses the term 'transcendental' to refer to inquiry into the possibility of experience, and because he thinks a 'deduction' of the categories is to be derived not from legal statutes but from showing how their instantiation serves as a necessary condition for the possibility of experience, he calls the task of justifying our use of the categories their 'transcendental deduction'. That is, his aim regarding the categories is to establish that they may be legitimately applied to objects of experience. To do this is to answer the skepticism that threatens the empirical claims we make using concepts such as *substance* and *cause*.

Synthesis and apperception

It might be thought that the course of our experience could never be anything but wholly contingent, and that nothing could possibly be known

about it *a priori*. For what appears to us in each moment is an existence entirely distinct from what appears in any other, and thus from what is occurring now, or has occurred in the past, nothing whatever could be inferred about what will happen in future moments, or even in the very next moment. This is in effect Hume's argument for skeptical doubt about both the basis in reason for our claims about future matters of fact and about causality or necessary connection.[1] Yet however obvious and irrefutable the Humean argument just given may appear, Kant's transcendental argument may be regarded as a direct assault on it.

The first point to make is that not just anything could count as "experience." Kant's argument in effect begins with a certain minimal conception of experience which he thinks even the most extreme skeptic must grant. Experience is something *manifold* through time, a *succession* of distinguishable contents that are present to a *subject* of that experience – and present to *numerically the same* subject throughout the time in which they appear. Unless we grant at least this much is true about our experience, we cannot even *raise skeptical questions* about whether the succession represents a continuing existence (such as a substance), still less an existence (such as a material object) distinct from the experienced contents themselves, or whether the successive occurrences are bound one to the next by causal relations. Kant's strategy is to begin with what we may call this indisputable conception of experience – to which we may give the name 'minimal experience', and then argue that minimal experience itself is possible only if its contents bear certain necessary relations to one another that may be known to obtain *a priori*.

The first stage of the argument points out that for there to be minimal experience, it must be possible for the subject to *apprehend* the series of experienced contents through an interval of time *as a series*, and then at the end of the interval to refer these contents to itself as to their self-identical subject. In order to represent what is manifold in an experience, Kant says, the subject must "first run through and then to take together this manifoldness, which action I call the **synthesis of apprehension**" (KrV A99).

Consider any successive experience – for example, the purely mental series constituted by someone's thinking a line of poetry, such as the first line of T. S. Eliot's *The Wasteland*:

> April is the cruellest month

This experience is made up of a series (or manifold) of distinct items spread out in time. We might distinguish the items in various ways – as phonemes, or syllables, or words. Let us consider the items to be words (thought or spoken silently by someone entertaining the line of poetry as a mere series of subjective representations). At the end of the time

interval in which I think this line to myself, I have before me as present the word 'month', but I also apprehend this word as having just been successively preceded by four no longer present words: 'April', 'is', 'the', and 'cruellest', these four words having previously occurred at four distinct times earlier in the series. Further, I must in this fifth moment in time be able to represent these four earlier contents as having just occurred, in that order, in four successive times. This means that at the end of the interval I must be able to *reproduce* these remembered contents as having previously occurred during the interval.

But not just any succession of contents could admit of this synthesis of reproduction. For our capacity to reproduce these contents depends on empirical laws, such as the Humean empirical laws of association by resemblance, contiguity, and cause and effect. Any "law of reproduction," however, "presupposes that appearances themselves are actually subject to such a rule . . . If cinnabar were now red, now black, now light, now heavy, . . . then my empirical imagination would never get the opportunity to think of heavy cinnabar on the occasion of the representation of the color red" (KrV A100–101). To go back to our example, there must be something in the *content* of the representations 'April', 'is', and so forth, that enables me to reproduce them at the end of the interval as having occurred, and as having occurred successively in the order in which I remember them. The associations in this line of poetry rest on an empirical synthesis, but even the succession of moments of time (or the successive survey of points in space), which are pure intuitions, require an analogous synthesis if their reproduction is to be possible. This pure synthesis, Kant argues, is the foundation of the empirical one – as the synthesis of reproduction of the moments of time at which we think (or silently speak) the words 'April', 'is', and so on is a condition of the possibility of reproducing those words in imagination. For even minimal experience to be possible, then, there must be something in the contents themselves, a combination among them and between them, that makes possible their orderly reproduction: Kant calls this the 'synthesis of reproduction'.

But now Kant argues for a third step (or third 'synthesis') that is necessary if even minimal experience is to be possible. It is not enough that I be able to *reproduce* the contents 'April', 'is', and so forth. I must also be able to *recognize* the reproduced contents as the *same* content-types that occurred earlier in the series. Otherwise, as Kant says, "all reproduction in the series of representations would be in vain. For it would be a new representation in our current state, which would not belong at all to the act through which it had been gradually generated, and its manifold would never constitute a whole, since it would lack the unity that only consciousness can obtain for it" (KrV A203). This means that I must be able to bring the original content and the reproduced content

under a common *concept* (for instance, the concept of the word-type 'April'). The contents themselves must be such that they are *conceptualizable* in determinate ways permitting their recognition under those concepts that make their reproduction possible – which means they must be combined already through what Kant calls the 'synthesis of recognition' (KrV A103).

The combination or synthesis of experienced contents that makes minimal experience possible is a certain relationship (or complex set of relationships) between the contents themselves. It is, namely, that set of relationships that makes us able to apprehend, reproduce, and conceptualize them. But Kant argues that these are not relationships that just happen to obtain between these contents, nor are they the kind of thing that could be merely contingently 'given' through the senses. For nothing can be combined for us in experience unless we combine it for ourselves through the self-activity of our understanding (KrV B130).

The synthesis on which possibility of experience depends, therefore, comes not from sensible data themselves, but rather on the *exercise of our faculties* on these data. This synthesis is therefore not a contingent and empirical but a necessary and *a priori* feature of experience. The fundamental synthesis that makes experience possible is the combination that enables us to attribute all our experiences to the same self-identical subject or 'I'. Kant uses the term 'apperception' to refer to self-awareness, and therefore he calls this fundamental synthesis the 'synthetic unity of apperception'.

"The 'I think' must be able to accompany all my representations; for otherwise something would be represented in me that could not be thought at all, which is as much as to say the representation would be impossible or else at least would be nothing for me" (KrV B131–132). Kant does *not* hold that all our representations are actually accompanied by the self-attribution 'I think'; he even follows Leibniz (and opposes Descartes and Locke) in thinking that most of our representations (or mental states) are *unconscious* (VA 7:135–136). But he argues that to call them *my* representations at all is to presuppose that they stand in a relationship to all my other representations, a relationship constituted by the activity of my understanding, that makes them in principle retrievable for my consciousness and ascribable by me to myself – without which they would be "nothing for me" and not elements of *my* experience (or *my* mental life) at all.

Objectivity and judgment

So far Kant's argument has established only a very abstract conclusion: that for even minimal experience to be possible, the contents of experience must be constrained and orderly in certain ways, and these ways are

determined not by what is given to the senses but by the self-activity of our understanding in constituting the unity of apperception – the fact that all my possible experiences are so connected that they are in principle ascribable to the same self-identical 'I'. The next step in Kant's argument is to identify this necessary combination itself with a specific concept that plays a fundamental role in our experience – namely, the concept of an *object* – and to identify the fundamental cognitive function that gives expression to necessary synthesis with the same function that led us to the categories – namely, the function of *judging* about objects.

In relation to what we have called 'minimal experience', objects have often been thought of only as something 'outside' that experience that might somehow "enter" it by being the *cause* of its contents. Kant himself thinks in these terms when he treats sensible intuition as the effect of an object on us. But in the Transcendental Deduction, his approach is different, and even revolutionary. For he wants to show us that owing to the necessary synthesis that makes even minimal experience possible, there must also occur in our experience something that plays the role of an object of the representations of minimal experience. In other words, his claim is that *minimal experience alone is not possible at all.* It is not possible because in order for even minimal experience to be possible, the merely subjective representations of minimal experience must stand in relation to objects that count as going beyond those merely subjective representations and laying claim to a kind of validity for any consciousness capable of experiencing at all.

We have seen that minimal experience is possible only if it is combined in such a way that its contents can be conceptualized. In relation to this, Kant says that an *object* is "that in the concept of which the manifold of a given intuition is **united**" [KrV B137]. Or in other words, an object is that which falls under those specific concepts in terms of which a synthesis necessary for minimal experience is thought. But how do we think such a necessary synthesis? Kant holds that our thought of such a synthesis takes the general form of a *judgment.* Every judgment grasps the *subject* of the judgment in thought and applies to it a *predicate,* which is a *concept.* For example, something given in our experience is judged to fall under the concept 'heavy' or 'red'; then the judgment may take the form: 'This is heavy' or 'This is red'. Or the subject itself may also be brought under a concept, such as 'piece of cinnabar', so that the judgment is 'This piece of cinnabar is heavy' or 'This piece of cinnabar is red'.

We are used to the picture according to which 'experience' is only minimal experience. In this picture, 'objects' are something entirely outside this experience, and 'judgments' are simply our attempt to "say something true" about these independent things. For this reason, Kant's

way of regarding things is bound to seem unfamiliar and puzzling. Why should we think of what we call 'objects' and 'judgments' as playing the role within experience that he assigns them?

To begin with, whatever else judgments may do, they obviously play the role of unifying our experience and representing to ourselves the ways different experiences are combined and organized. When I judge that the object I hold in my hand is red and heavy, I bring my present perceptions under general concepts, relating what I now see and feel to other things I count as 'red' and 'heavy'. When I judge that cinnabar is red and heavy, then I bring an indefinite number of possible perceptions (in which I might perceive pieces of cinnabar) under these same concepts. Moreover, judgments represent a way of organizing what is given in experience that has reference to features of it that are in some sense necessary rather than contingent. And it is this necessity that grounds the *objectivity* of judgment, the fact that a true judgment is valid not only for the subject who judges, but for all subjects who might consider the judgment. This is because the orderedness among appearances expressed in the judgment is due not merely to contingent features of the way they are given in sensation, but to the activity of our understanding in synthesizing them so as to make experience possible. This makes the ground of the objectivity of judgments *a priori*, and guarantees the validity of a true judgment for any subject of experience who judges at all.

Kant thinks that the objectivity of a judgment, or its universal validity, is what we express through the copula 'is' that links subject to predicate: "For this word designates the relation of the representations to the original apperception and its **necessary unity**, even if the judgment itself is empirical, hence contingent, e.g. 'Bodies are heavy'. By that, to be sure, I do not mean to say that these representations **necessarily** belong **to one another** in the empirical intuition, but rather that they belong to one another **in virtue of the necessary unity** of the apperception in the synthesis of intuitions" (KrV B142). Thus Kant's position is *not* that in judging this piece of cinnabar to be red, I am judging that *it* is *necessarily* red. For it is only a contingent, empirical fact that cinnabar is red. Rather, the point is that my judgment that it is (contingently) red carries with it a certain normative force, not only for me at this time, but for me at other times, and indeed for any possible subject. For if my judgment is true, then it will be binding on me also at other times to judge that this piece of cinnabar is red, and this same truth is also binding on any other possible subjects who might judge about the color of this sample of cinnabar.

Kant thinks this peculiar normative necessity of true judgments, their universal applicability, is to be understood transcendentally as the way in which their truth supplies one of those synthetic connections among experiences that are necessary for experience to be possible at all. In other words, the normative force of this empirical judgment lies in the

fact that it is an instance of the kind of necessary connection among the contents of experience that must obtain if even minimal experience is to be possible. The necessity of the syntheses expressed in judgments about cinnabar as an object of judgment is what makes the true judgment, even when its truth is contingent, valid equally for all possible perceivers of this cinnabar. In other words, it is what makes a true judgment *objectively* valid, and therefore what makes the piece of cinnabar itself into an *object*, rather than merely a contingent collection of representations present to a single isolated subject.

The traditional way to think about objectivity – about the validity for all subjects of cognition belonging to true judgments about objects – is to see it as parasitic on the existence "in themselves" – apart from any possible consciousness – of the things about which we are judging, and their objective properties. It is a consequence of Kant's "Copernican revolution" that we must revise our way of thinking and treat the universal validity for consciousness in general as constituting the objectivity of objects.

The synthetic unity of apperception (or the necessary relation of subjective representations in minimal experience) therefore guarantees that any possible experience is more than minimal experience, in that any possible experience must be conceptualized in such a way that it is judged to contain objects about which the subject of that experience can make judgments that are true – judgments that are valid for all possible subjects and may therefore be regarded as 'corresponding' to the object the judgment is about.

Some philosophers have thought that Kant's revolutionary transcendental way of regarding experience involves a denial of the "correspondence" theory of truth. Taken literally, this is clearly mistaken, since Kant explicitly affirms, as a "nominal" definition, that truth consists in correspondence of a judgment to an object, and he denies that any "real" definition of truth is possible (KrV A57–59/B82–83).[2] That is to say, the correspondence theory tells us what we mean by 'truth' but there cannot be any account of truth that we could use as a criterion in deciding what is true and what is not. There may be something correct in the thought that Kant denies the correspondence theory if it means to say that from the standpoint of the role played by pure *understanding* and its *conceptualization* in constituting experience, he treats 'correspondence' not as some mysterious *sui generis* relation between a mental act and some thing wholly alien to what is experienced, but rather as a relation *constituted* by the way that the items falling under certain concepts play a necessary role in satisfying the transcendental conditions under which the experience of any and all subjects becomes possible.

However, the thought that Kant is rejecting the correspondence theory of truth is also false at a deeper level, since it ignores the necessary role

of *intuition* in constituting the objectivity of the things to which our judgments correspond. It is crucial for Kant that what is conceptualized as playing the role of an object in unifying a manifold of intuition should also be grasped as something *given in* intuition (if not directly, then by way of some intuition about another object that is connected in a rule-governed way with the object we are judging about). Hence for Kant the judgments we make about an object, if they are true, should *correspond* to the object as to a thing that is, or at least that could be, somehow immediately present to us in space and time. The judgment that all samples of cinnabar are heavy may of course be made by someone who has never lifted a piece of cinnabar. But the truth, and even the significance, of this judgment depends on the fact that the concept 'cinnabar' represents a certain synthesis of possible experiences for any subject that are valid for all subjects and that could in principle be attested in the intuition of any given subject – as when that subject comes across a sample of cinnabar and feels its heft in her own hands. It is important then to recognize that the truth of the judgment is not merely present as an abstract transcendental necessity required for the synthetic unity of experience, but is at the same time something we intuit as given to our senses, and therefore as an *intuited* fact *corresponding* to our judgment.

II objects, categories, and schemata

In the "metaphysical deduction" Kant claims that the twelve categories correspond to the forms of our judgment, so that these twelve concepts are possessed by any subject who has the capacity to judge at all. What is crucial for Kant's transcendental deduction of the categories, however, is not merely that they correspond to the fundamental forms of all the judgments that we may make, but also that they necessarily apply to any possible objects that may be given to us through sensible intuition (KrV B144–145). The key to taking this new step, in Kant's view, seems to be that space and time are given not merely as forms of intuition, but also as *a priori* intuitions that are themselves unities (KrV B160–161). Thus there is a correspondence between the *synthetic* unity that apperception brings to experience and the intuited unity of what is given in experience. Perhaps the point is best put by saying that space and time are given in intuition as a single unified framework or system of relations between objects and their changes, so that the objective unity of experience that is *thought* by the understanding in apperception and expressed in the universal validity of judgments can be regarded as the same unity of experience that is immediately given in intuition.

This provides Kant with a way of accounting, within a transcendental idealist framework, for the reality of natural objects that can never actu-

ally be sensed by subjects like ourselves (either because they are too small, too distant, or because they existed at a time in the world's history prior to the existence of any subjects like ourselves). Because experience constitutes a unified, law-governed system of objects in space and time, Kant specifies that anything is "actual" or "real" if it is connected with an actual sensation according to the laws of experience (KrV A218/B266). Thus dinosaurs count as real at the past time when they existed because actual sensations of fossilized dinosaur-bones can be connected according to determinate causal laws (those governing geological processes of fossilization, Carbon-14 decay, and so on) with the past existence of the animals of whose bones they are an empirical record.

But the apriority of the categories, combined with the claim that they necessarily apply to objects of experience, leaves us with a serious problem, in the form of a question that might seem unanswerable: How do we come to recognize instances of the categories (of quantity, reality, substance, cause, and so on) within our experience in particular cases? If these were empirical concepts, such as 'dog' or 'water' or 'charcoal', the question how we recognize instances of them would virtually answer itself. For since we acquire empirical concepts through sensible intuition of their instances (or, in some cases, of the instances of partial concepts of which they are composed), the same experiences that afford us these concepts also afford us experiences in which their instances are recognized. But our possession of *a priori* concepts is not to be accounted for in this way – in fact, Kant's transcendental deduction has argued that our possession and even our use of them is more a matter of what our understanding actively brings to experience than of what is given to us in it. If causes are in our experience because we have to apply the concept of cause in order to make judgments and thereby to constitute experience as a unity for our apperception, why should we apply the concept to any one object in our experience more than another, and how can we hope to distinguish cases of causal connection in the empirical world from cases in which there is no causal connection?

Kant proposes to answer this question in the next chapter after the Transcendental Deduction, which he entitles "On the Schematism of the Pure Concepts of Understanding." For Kant the "schema" of a concept is a condition of sensibility under which the concept can be applied to an object (KrV A140/B179). Kant compares schemata to images and says that the schematism of the categories is accomplished by the faculty of "productive imagination" (KrV A141–142/B180–181). We can think of ourselves as applying the concept 'dog' through the fact that we have a sensible image of a dog (or rather, probably an indefinitely large family of such images), reproduced from past experiences of dogs, with which we can compare a given object in sensation to determine whether

it falls under our concept 'dog'. Analogously, we are invited to think of the schema of a concept as a representation created *a priori* by our imagination. But it leads us into perplexity to think of the recognition of instances of concepts (whether pure or *a priori*) as being accomplished by means of representations that are like mental images. For one thing, when Kant describes the schema as a representation, he has to describe it, paradoxically, as one that is "homogeneous" with a concept and with a sensation (KrV A137/B176), or at the same time intellectual and sensible (KrV A138/B177). More basically, the problem is that the ability to recognize instances of a concept always remains an *ability*, and can never be simply identified with having a mental representation of any kind. For whatever picture of a dog (or a cause) we might have in our minds, we recognize dogs (or causes) by means of it only if we possess the *ability* to apply the picture *correctly* to whatever comes before us in experience.

A "schema" is described as a "representation of a general procedure of the imagination for providing a concept with its image" (KrV A140/B179–180). Schemata and concepts are both *rules* governing the proper exercise of our mental capacities. A concept is a rule for combining other representations under a common representation. Schemata are rules for displaying or recognizing instances of a concept in sensible intuition. Kant denies, however, that there is any image corresponding to a pure concept of the understanding (KrV A142/B181). So when he comes to characterize the schemata of the categories, what he gives us is neither a mental image nor an ability but rather a kind of *recognizable pattern* within experience, and in particular within intervals or sequences in time. We are to interpret these general descriptions of patterns in experience as rules or procedures for identifying intuitable instances of empirical concepts that fall under the schema.

The schema of reality is "a concept of the understanding to which a sensation in general corresponds, that therefore, the concept of which in itself indicates a being (in time)" (KrV A143/B182). The schema of substance is "persistence of the real in time," while the schema of causality is "the real upon which, whenever it is posited, something else always follows" (KrV A144/B183). These formulas do not describe anything like mental images; having the schema of a category is having the ability to recognize certain very abstractly characterized occurrences in time, especially in the context of making judgments. To be able to apply the concept of cause, for instance, is to be able to judge that something real has occurred upon which something else always follows (according to a lawlike regularity). If this is right, then perhaps Kant is not so far off in ascribing schematism of the understanding to the productive imagination, since it is not unreasonable to think that the ability to make judgments of this sort is closely allied to the ability to form images. And if

someone were to wonder how it is possible that we should have such an ability, Kant will answer such a skeptic in the next part of the Transcendental Analytic, where he shows that the very possibility of experience depends on our ability to make such judgments.

III the principles of pure understanding

The final phase, and the culmination, of Kant's long, interdependent chain of argument in the Transcendental Analytic attempts to vindicate *a priori* a certain conception of nature, the one that grounds the methods and procedures of modern mathematical (and more specifically, Newtonian) physics. The picture of nature is that of a system of relations in space and time constituted by a single material substance, extended through space and constantly altering in time, whose quantity remains constant through all natural change. The system may also be treated as containing a plurality of substances, since the spatial parts of material substance can be treated as distinct existences. The states of these different substances are mutually determined at any given instant by causal relationships between the substances, and in which the perpetual changes in the states of substances follow necessary causal regularities. Both the formal relationships obtaining in the system of nature (the relationships in space and time) and the material relationships (depending on the reality of substance and the causal powers of the states of different substances) are quantities, and the causal laws determining them are therefore essentially mathematical in form.

The "Principles" chapter is divided into four subsections, corresponding to the four groups of categories. The first, "Axioms of Intuition," attempts to establish the application of concepts of quantity to the "extensive magnitudes" of space and time, thereby making mathematics applicable to spatio-temporal objects (KrV A161–165/B202–207). The second, "Anticipations of Perception" (KrV A165–176/B207–218), does the same regarding "intensive magnitudes", such as the causal powers of objects. The fourth, "Postulates of Empirical Thinking" (KrV A218–235/B265–287), specifies the application of the modal categories (possibility, actuality, necessity) – which, Kant says, never concern objects as such, but only the relation of our understanding to them. The most important subsection, and the one we will discuss here, is the "Analogies of Experience" (KrV A176–218/B218–265). It deals with the categories of relation, and argues that all change in the empirical world involves certain necessary connections. The First Analogy (KrV A182–A189/B224–232) says that every change is an alteration (or change of states) of a single persisting *substance* (matter), whose quantity never increases or diminishes. The Second Analogy (KrV A189–211/B232–256) holds that

all these alterations follow necessary (causal) laws, determining that the state that is prior in time is necessarily succeeded by the state that follows it. The Third Analogy (KrV A211–215/B256–262) says that at any given time, there exist simultaneous causal relations mutually determining the states of different substances (that is, of the spatial parts of the one persisting substance). In this way, the Analogies are intended to ground *a priori* the modern scientific conception of the world as a system of material objects governed by deterministic causal laws.

The conditions of time-determination

The basic idea behind Kant's arguments for the Analogies is that our experience is determined in time. That is, our experiences occur through intervals of time, and there is a succession in time within the manifold of experience. Further, there is a fact of the matter about whether (and how much) time has passed, and for any two happenings in time, there is a fact of the matter about whether they happened at the same time, or (if they did not) about which of them preceded and which one followed the other. The thesis that our experience is (in this sense) "determined in time" may be regarded as part of the conception of 'minimal experience' from which Kant argued in the Transcendental Deduction. For minimal experience is the occurrence of a succession of representations through time, which is possible only if there is a fact of the matter about the passage of time and about the succession of events in time. Further, if we admit one important conclusion of the Transcendental Deduction, that minimal experience is possible only if experience also contains *objects* (in addition to subjective representations), then we must also take the states of these objects to be determined in time.

Time-determination, in this sense, means that there is a *fact of the matter* about the duration of states and about the order of their succession. It does *not* directly entail that there is anything (such as a substance) that endures through intervals of time, or that successive states are causally determined by what preceded them. In order to establish those conclusions (as he intends to do in the Analogies), Kant must appeal in addition to further premises having to do with the conditions under which we can regard our experience as determined in time.

One of these is that if there is a fact of the matter about time-determination regarding duration, succession, and simultaneity, then it is a fact that is *knowable by us*, at least in principle. This premise may be regarded as something we are justified in supposing if, as Kant does, we take space, time, and the objects of experience to be not things as they exist in themselves but things as they appear to us, things considered insofar as they fall under the conditions of our sensible intuition (of which time itself is only a condition). To suppose about objects that

their states might be determined in time, but that these determinations could not be knowable by us even in principle, is in effect to treat the facts of time-determination as lying beyond the reach of our cognitive faculties, and this contradicts the assumption that they are facts precisely *for* those faculties. (It is reasonable for Kant to presuppose transcendental idealism in the Analogies, because transcendental idealism has already been argued for in the Transcendental Aesthetic, whose results are being taken for granted even in the way Kant is raising the problems the Analogies are trying to solve.)

The other crucial premise Kant needs in the Analogies is one that he often expresses (perplexingly) by saying that "time cannot be perceived in itself" (KrV A176/B219, B225, B233, B257). It means that facts about the determination of time cannot be directly known or read off our experience just in virtue of the fact that our experience itself is temporal. There is in principle always a distinction, in other words, between one time interval's *seeming* (subjectively) longer than another and its actually *being* longer – it is not known to be longer just because it seems so to us. Likewise, the fact that we subjectively experience one state as happening earlier than another does not entail that it really does happen earlier. (In dreams, for instance, sometimes what comes later in the narrative sequence of the dream – hence what seems later – is something we later know was occasioned by an external event – the flashing of a light or banging of a door in the room where we are sleeping – which we know happened at the start of the dream.) This entails that the objective determination of time must depend on (in principle knowable) facts about the objects and objective occurrences with which we are presented in time. In other words, the time-determination that is required if (even 'minimal') experience is to be possible can exist only if there are certain necessary connections between the objective occurrences taking place in the objective world of appearances.

The three Analogies, and the Refutation of Idealism

If time cannot be perceived in itself, then the duration of time must be represented in experience by something persisting. According to the First Analogy, this persisting thing is material substance and all change in experience must consist in the alteration of its states. If there is to be a fact of the matter about which states come first and which follow, then this can consist only in a necessary rule governing the succession of states, and determining that the one must precede, and the other must follow upon it. According to the Second Analogy, this necessary relation between successive states is a causal law. If the states of different substances (the one substance distinguished into many substances by their spatial locations) are to be objectively determined at any given time as

simultaneous, then there must be a necessary rule that connects each state with the other. According to the Third Analogy, this rule is a causal law determining a reciprocity or "community" between these states. (The prototype of such a law for Kant is obviously the Newtonian law of gravitation, according to which at every moment every pair of bodies exert a force on one another according to a law determined only by their masses and the distance between them.)

The nature of substance and the causal laws determining its states concerns, as Kant says, the "existence" of appearances rather than merely their perception. That is, the nature of substance and causal laws are objective facts about the world which, if we knew them, would amount to a knowledge of the determination of time (facts about the duration of intervals and the simultaneity or succession of moments). We obviously do not make everyday judgments about time by means of our knowledge of these fundamental necessary connections in the physical world, but Kant's argument is that we must presuppose such connections in order to be assured that there are such facts to be known.

Closely related to the Analogies of Experience is an argument present only in the second edition of the *Critique*, the "Refutation of Idealism" (KrV B274–279). The "idealism" Kant sets out to refute is not, of course, the "transcendental" or "critical" idealism on which he builds his account of how synthetic *a priori* cognition is possible. It is rather the position that takes us to know the existence of objects external to us in space only through a causal inference from our subjective perceptions, and therefore either denies (as did Berkeley) or doubts (as did Descartes, until his Sixth Meditation) the existence of such objects. Kant's counterclaim is that our awareness of external objects is immediate, an awareness through intuition, and not by any sort of causal inference (KrV B276). In the first edition of the *Critique,* Kant seems content to establish this conclusion merely by appealing to the fact that space is a form of outer intuition, in which empirical objects are immediately given. But partly in response to the interpretation of his own position as a version of Berkeleyan idealism, in the second edition he sought a stronger argument that would establish the conclusion that the objects in space we immediately intuit must be *distinct from all our representations*, and that such distinct objects must be presupposed if there is to be any determination in time even of the sequence of our subjective representations themselves.

Kant's argument depends on taking seriously the idea that the subjective representations through which we acquire a perspectival awareness of the world are never more than seemings. They provide us with indispensable cognitive access to the objective world, but they themselves provide us with no sort of infallible knowledge of that world, nor do they even constitute part of that world. Our subjective thoughts are

not, as Descartes held, objectively existing modes inhering in a thinking substance (as shapes and motions inhere in extended substances). Inner sense provides us with no immediate and infallible access to an "inner" reality, from which all facts about "outer" reality must be inferred if we are to know them at all. On the contrary, the only reality of which we have immediate intuition is the external reality of material substances whose time-determination depends on their duration and the law-governed succession of their states. Hence if there is to be an objective fact even about the time-determination of our subjective representations, then this fact must be derivative from the temporal facts about the states of external substances that are truly objective and distinct from these representations, which are nothing but our perspectival mode of access to the objective material world of real things in space.

Skepticism about the external world depends on the picture of an "inner world" of representations from which we would have somehow to "advance" (by causal inferences) if we are to arrive at an "external" world. It depends on calling into question the means of this supposed "advance," thus leaving us with the possibility that there is no reality at all except the "inner" reality. Kant describes his "Refutation of Idealism" as "turning the tables" on this skepticism, by showing that the skeptic's picture itself is incoherent, that even the time-determination the skeptic needs in order to postulate a succession of "inner" states is parasitic on there being an external objective reality to which our subjective states provide us access.

The skeptic's picture also depends on assuming that "objective reality" consists in the separate existence of an order of things that is entirely distinct from what is given in our experience, so that our beliefs about it would have to be justified by some sort of inference from what is given within our experience to what is present in that independent order. Kant's transcendental idealism is meant to undercut that picture also, by showing that "objectivity" is constituted instead by that order of what is given in our experience (an order of substances in space and time, whose simultaneous states are related by causal reciprocity and whose successive states are governed by causal laws) – an order which is shown to be necessary if experience is to be possible at all. But Kant's transcendental idealism itself is a new picture that has seemed to many to be problematic, or even internally incoherent. Our next task will be to explore those problems and see if a resolution to them is possible.

further reading

Eckart Förster (ed.), *Kant's Transcendental Deductions*. Stanford: Stanford University Press, 1989.

Dieter Henrich, *The Unity of Reason*, ed. R. Velkley, tr. J. Edwards. Cambridge, MA: Harvard University Press, 1994.
Pierre Keller, *Kant and the Demands of Self-Consciousness*. New York: Cambridge University Press, 1998.
Patricia Kitcher, *Kant's Transcendental Psychology*. New York: Oxford University Press, 1990.
Béatrice Longuenesse, *Kant on the Human Standpoint*. Cambridge, England: Cambridge University Press, 2005.
Arthur Melnick, *Kant's Analogies of Experience*. Chicago: University of Chicago Press, 1973.

notes

1 See Hume, *Enquiry Concerning Human Understanding*, ed. Eric Steinberg (Indianapolis: Hackett, 1977), pp. 18–19.

2 Perhaps the best-known example of this thought is found in Hilary Putnam, *Reason, Truth and History* (Cambridge, England: Cambridge University Press, 1981), pp. 60–4. Putnam of course realizes that Kant explicitly affirms correspondence as a nominal definition of truth, but tries to "extract" from Kant's texts the thought that a true statement is whatever a rational being would accept on sufficient evidence. But this looks like an attempt at a 'real' definition of truth, and Kant denies that any real definition of truth can ever be given. Putnam also confuses the issue by claiming that a correspondence theorist must believe our thoughts or statements correspond to "things in themselves," and then inferring from Kant's denial that we can know things in themselves that he must be denying a correspondence theory of truth. Some correspondence theorists no doubt take truth to be correspondence to things in themselves, but in the notion of 'correspondence to reality' there is no commitment to what kinds of real things we have to mean. In Kant's nominal definition of truth, it obviously means nothing but the correspondence of judgments about appearances or empirical objects to those same appearances or empirical objects. If Putnam means to suggest that for Kant appearances or empirical objects are not "realities," and that only things in themselves are real, then he clearly misunderstands Kant on this point too.

chapter 4

the limits of cognition and the ideas of reason

I transcendental idealism

Kant builds his critique of reason on a novel doctrine about the nature of human knowledge and of its objects, to which he gives the name 'transcendental idealism' or 'critical idealism'. The doctrine can even be stated with apparent simplicity: We can have cognition of appearances but not of things in themselves. But it is far from clear what the doctrine means, and especially unclear what sort of restriction it is supposed to place on our knowledge. Some readers of Kant have seen the restriction as trivial, so trivial as to be utterly meaningless, even bordering on incoherence. They have criticized Kant not for denying that we can know 'things in themselves' but rather for thinking that the notion of a 'thing in itself' even makes sense. If by a 'thing in itself' we mean a thing standing outside any relation to our cognitive powers, then of course it seems impossible for us to know such things; perhaps it is even self-contradictory to suppose that we could so much as think of them. Other readers have seen transcendental idealism as a radical departure from common sense, a form of skepticism at least as extreme as any Kant might have been trying to combat. To them it seems that Kant is trying (like Berkeley) to reduce all objects of our knowledge to mere ghostly representations in our minds. He is denying us the capacity to know anything whatever about any genuine (that is, any extra-mental) reality.

Some of the perplexities about transcendental idealism are due to people's inability or unwillingness to see the philosophical problems of knowledge as Kant sees them, and therefore to consider seriously his proposed solution to them. But I think much of the puzzlement about transcendental idealism arises from the fact that Kant himself formulates transcendental idealism in a variety of ways, and it is not at all clear how, or whether, his statements of it can all be reconciled, or taken as statements of a single, self-consistent doctrine. I think Kant's central

formulations suggest two quite distinct and mutually incompatible doctrines. My first aim will be to describe these two doctrines. After that I will argue *first*, that we cannot choose between them as interpretations of his meaning on merely textual grounds (since both are undeniably present in his texts), and *second*, that one interpretation is clearly preferable to the other, though entirely on non-textual (that is, philosophical) grounds.

The causality interpretation

Kant often distinguishes appearances from things in themselves through locutions like the following: "What the objects may be in themselves would still never be known through the most enlightened cognition of their appearance, which alone is given to us" (KrV A43/B60). "Objects in themselves are not known to us at all, and what we call external objects are nothing other than mere representations of our sensibility, whose form is space, but whose true correlate, i.e. the thing in itself, is not and cannot be cognized through them" (KrV A30/B45). Passages like these suggest that things existing in themselves are entities distinct from '*their* appearances' – which are subjective states caused in us by these things. Real things (things in themselves) cause appearances. Appearances have no existence in themselves, being only representations in us. "Appearances do not exist in themselves, but only relative to the [subject] insofar as it has senses" (KrV B164). "But we should consider that bodies are not objects in themselves that are present to us, but rather a mere appearance of who knows what unknown object; that motion is not the effect of this unknown cause, but merely the appearance of its influence on our senses; that consequently neither of these is something outside us, but both are merely representations in us" (KrV A387). We may call the version of transcendental idealism that follows this picture the 'causality interpretation', because its fundamental point is that the relationship between things in themselves and appearances is a causal relation: appearances are subjective states in us, that are *caused* by things in themselves outside us. Kant seldom uses the term 'cause' to describe the relation of things in themselves to appearances, but he frequently uses the term 'ground' – perhaps because it seems to him more abstract and metaphysically non-committal, better suited to express a relation that can never be cognized empirically but only thought through pure understanding. It might also be called simply the 'non-identity' interpretation, since the point will be that relations like cause and ground require different entities as their relata. If a given appearance – say this chair – is grounded on or caused by some thing in itself, then at the very least, it cannot be identical with that very thing that grounds or causes it; so it has to be a different thing.[1]

The identity interpretation

In other passages, transcendental idealism is formulated so as to present us with a very different picture. "We can have cognition of no object as a thing in itself, but only insofar as it is an object of sensible intuition, i.e. as an appearance . . . We assume the distinction between things as objects of appearance and the very same things as things in themselves, which our critique has made necessary" (KrV B xxvi–xxvii). Here Kant does not distinguish between two separate entities, but rather between the same entity as it appears (considered in its relation to our cognitive faculties) and as it exists in itself (considered apart from that relation). We may call this the 'identity interpretation' because its fundamental point is that every appearance is *identical* to a thing in itself, and the distinction is not between two different entities but between two ways of thinking about or referring to the same entity. "The things we intuit are not in themselves what we intuit them to be, nor are their relations so constituted in themselves as they appear to us" (KrV A42/B59).

> The concept of appearances . . . already . . . justifies the division of objects into *phenomena* and *noumena*, thus also the division of the world into a world of the senses and of the understanding . . . For if the senses merely represent something to us **as it appears**, then this something must also be in itself a thing, an object of a non-sensible intuition, i.e. of the understanding . . . through which, namely, objects are represented to us **as they are**, in contrast to the empirical use of our understanding, in which things are only cognized **as they appear**. (KrV A249)

On the identity interpretation, appearances are not merely subjective entities or states of our minds; they do have an existence in themselves. The force of transcendental idealism is not to demote them, so to speak, from reality to ideality, but rather to limit our cognition of real entities to those features of them that stand in determinate relations to our cognitive faculties. Some things in themselves may not be intuitable by us, and so may not be appearances. But every appearance also has an existence in itself. The identity interpretation is also called the "two conceptions" interpretation, because it holds that appearances are not distinct entities from things in themselves, but the same entities, conceived or referred to in different ways. To call an entity an 'appearance' is to refer to it as something intuited by us and standing in relation to our faculties; to call it a thing in itself is to refer to it as it exists apart from that relation.

The causality interpretation is also sometimes called the 'two worlds' interpretation, because it holds that appearances and things in themselves constitute two different worlds, two separate realms of distinct entities. Yet the very same entities could belong to two different worlds,

just as exactly the same person might hold dual citizenship or belong to two different clubs. Likewise, there is nothing to prevent someone from saying that the relation between a thing in itself and the corresponding appearance is one of identity and yet distinguishing between sensible and intelligible worlds, to both of which these self-identical things belong. Thus the identity interpretation has just as much right as the causality interpretation to call itself a 'two worlds' interpretation. According to it, something belongs to the world of sense insofar as it is an object of our sensible intuition; it belongs to the world of the pure understanding insofar as we abstract from this and consider it through pure understanding not as we sensibly intuit it, but as it exists independently of our capacity sensibly to intuit it. The crucial thing about the causality interpretation of transcendental idealism is that it holds that no individual appearance or phenomenon is identical to any thing existing in itself or noumenon, and no noumenon or thing existing in itself is identical to any phenomenon or appearance.

Points in common, points of difference

Kant may have shifted between two ways of expressing his doctrine because the two interpretations agree on the four points that are most important to him. These are:

1 Real things exist.
2 They cause representations in us.
3 Objects of our cognition are given to us through the senses and thought through the understanding.
4 Sensing and thinking are subject to conditions that make synthetic *a priori* cognition possible.

Yet the two interpretations appear to yield different (incompatible) answers to the following three questions:

1 Is an appearance the very same entity as a thing in itself? The causality interpretation says no, the identity interpretation says yes.
2 Are appearances *caused* by things in themselves? The causality interpretation says yes, the identity interpretation says no.
3 Do the bodies we cognize have an existence in themselves? The causality interpretation says no, the identity interpretation says yes.

Some interpreters of Kant, when they become aware of these divergences, respond by saying that there is no significant difference between the two interpretations, that they are only "two ways of saying the same thing."[2] These interpreters are probably faithful to Kant's intentions, since it looks as if he thought the two ways of talking about appearances and things in themselves are interchangeable and involve no difference in doctrine. But someone can intend to speak self-consistently and yet fail to do so; and it looks like this is what has happened to Kant in this

case. For no entity stands to *itself* in the relation of cause to effect. Transcendental idealism is no intelligible doctrine at all if it cannot give self-consistent answers to the above three questions.

Perhaps this is why interpreters who take this line sometimes disagree over *which* "same thing" Kant's different pronouncements are two ways of saying (some in effect taking the identity interpretation, others the causality interpretation).[3] Some hold that things in themselves *are* appearances, only they are being considered in abstraction from the conditions under which they appear. We speak of appearances as if they were effects of these things only because our intuition involves being affected by things. Others think in effect that Kant regards appearances as truly caused by things in themselves, but lets them count as the same things for the purposes of everyday discourse (as with Berkeley, there is one way of talking with the learned, another with the vulgar). But both sides ought to admit that in the end there must be only one correct way of parsing the Kantian double-talk, and the other way must be regarded as at most a *façon de parler* and not Kant's genuine doctrine. The question is: Which way is the genuine doctrine and which merely the *façon de parler*?

Kant occasionally tries to combine "causality interpretation" talk with "identity interpretation" talk. When he does, the result is simply nonsense and self-contradiction:

> I say that things as objects of our senses existing outside us are given, but we know nothing of what they may be in themselves, cognizing only their appearances, that is, the representations which they cause in us by affecting our senses. Consequently, I grant by all means that there are bodies outside us, that is, things which, though quite unknown to us as to what they are in themselves, we still cognize by the representations which their influence on our sensibility procures us, and which we call bodies, a term signifying merely the appearance of the thing which is unknown to us but not the less actual. (P 4:289)

The first sentence here says that objects of the senses are given to our cognition, but then denies that we cognize these objects, saying instead that we cognize an entirely different set of objects (different from the ones he has just said are given). The second sentence infers from this that there are bodies outside us, but proceeds to say that it is not these bodies (that is, the entities Kant has just introduced to us as 'bodies') that we call 'bodies', but rather *bodies* are a wholly different set of entities. Such Orwellian doubletalk seems to be the inevitable result of trying to combine the causality interpretation with the identity interpretation while supposing that they are just two ways of saying the same thing.

The only way I can see that we might avoid having to choose between the two interpretations is to say that the questions that separate them are unanswerable or meaningless questions, because it is improper even

to ask whether an appearance is the same entity as its corresponding thing in itself.[4] Peter Geach, for example, held that all meaningful assertions of identity are relative to some concept: Hesperus and Phosphorus are the same *planet*, Tully and Cicero are the same *human being*. But 'Geach' is the same word *type* as 'Geach', without being the same word *token*, and Heraclitus *can* step into the same *river* twice, though not into the same *flowing water*.[5] On Geach's view, it is illegitimate to predicate identity of two things without specifying the concept under which one is claiming they are identical, and it is also illegitimate to circumvent this requirement by employing for this purpose a 'dummy' concept such as 'thing', 'being', 'object', or 'entity'. Using Geach's view, it could be argued that there is no common concept under which we may bring a phenomenon or appearance and the corresponding noumenon or thing in itself, since the former sort of entity always falls under a concept with intuitive content, and the latter never can. Therefore, it is nonsensical either to assert or to deny that an appearance is the same entity as the corresponding thing in itself, and the issue separating the causality from the identity interpretation is a bogus one.

Kant never explicitly endorses any view of this kind, of course. If he did, he would surely express reservations about saying *either* that appearances are identical with things in themselves *or* that they are not identical. He would avoid locutions in which an appearance is said to be the "very same thing" as a thing in itself. But far from all this, he never even seems to be aware that there might be so much as a problem reconciling his talk of things in themselves "grounding" appearances with his talk of things "as they appear" and "as they are in themselves."

Kant treats identity as a "concept of reflection" – a concept involving a comparison of two ways of representing an object, even of ways of representing it through different faculties (KrV A260–265/B316–321). He holds that when objects are represented in pure understanding, the criterion to be used in individuating them is the Leibnizian one – the identity of indiscernibles – while when they are given to us through the senses, the principle of their individuation is their positions in space. These points might suggest to someone of Geachian sympathies that for Kant too there is no univocal criterion of identity or non-identity that can be used on appearances and the "corresponding" things in themselves, hence that strictly speaking, neither identity nor non-identity can be said to hold between them. The issue, it might be claimed, is a bogus one, because the conditions for applying either identity or non-identity would not hold when we compare appearances with things in themselves. For each concept would carry with it a different criterion for identity, and there would be no common criterion in terms of which identity could be either affirmed or denied across the conceptual gulf that separates appearances from things in themselves.

Yet if this were right, then one would have thought that instead of speaking casually in both ways (asserting identity and at the same time employing relations that entail non-identity) Kant would want to refrain from talking either way. Yet Kant shows no reluctance to use 'thing' and 'object' in ways forbidden by such a view and even directly to assert the identity of things in themselves with appearances: "We assume the distinction between things as objects of appearance and the very same things as things in themselves . . ." (KrV B xxvii). In any case, it seems gratuitously obscurantist to import the gross implausibility of Geach's view of identity (for we cannot deny that Geach's views about identity are pretty hard to swallow) into Kant merely in order to rescue his readers from facing up to a difficult exegetical dilemma.[6]

Besides, there is an obvious asymmetry between Kant's two criteria of identity that stands in the way of using his views about identity to get rid of the issue. Although things in themselves cannot be sensed, appearances can be thought through the pure understanding, simply by thinking of them in abstraction from the ways they can appear to us. Thus while the sensible criterion for identity cannot apply across the gulf separating phenomena from noumena, the intelligible criterion can apply. This seems in fact to be precisely the way in which Kant himself often arrives at the concept of a noumenon or thing in itself. We begin with particular sensible things (appearances) and then represent them as they are apart from our sensing them, solely through concepts of understanding (KrV A238/B298). It is true that we are thereby abstracting also from the criterion of their identity or distinctness *as appearances*. But the criterion of identity or distinctness involving things in themselves (if we need one at all) is Leibnizian (sameness or distinctness of concept). Once we have abstracted from the sensible – e.g. the spatio-temporal – properties of the object as appearance, hence from our empirical cognition of it, it must be the same as *itself* thought solely through pure concepts of the understanding and distinct from any thing which is represented as *other* than it (for instance, from a *different* appearance regarded as it is in itself). No doubt Bishop Berkeley would object to our act of abstraction; he would deny that we could think without contradiction a sensible thing as it might exist apart from being sensed. But this seems to be a point on which Kant and Berkeley evidently disagree.

Kant of course denies that we can ever have *cognition* of an object as it is in itself, because we can have no sensible intuition of it – as it is in itself. But he seems to regard it as entirely permissible and even inevitable that we should be able to *think* the phenomenal objects around us solely through pure concepts of the understanding, hence as they are in themselves. If I arrive at the concept of the chair in the corner first by cognizing it empirically and then by abstracting from those conditions of cognition, so that I think of it existing in itself outside those

conditions, then it is obvious that I am thinking of the same object, not of two different objects. It is also clear that when I think of it the second way, I am thinking of *it*, and not of its cause (if it has one). From this point of view, the causality interpretation seems utterly unmotivated and even nonsensical.

The problem arises, however, because Kant *also* wants to arrive at the concept of a thing existing in itself in another way. He starts from the fact that our empirical cognition results from the affection of our sensibility by something outside us. This leads him to think that there must be a cause acting on our sensibility from outside, making it possible for us to intuit appearances, which are then conceived as the effects of this cause. Of course it would be open to him to think of this for each case of sensible intuition as the appearance acting on our sensibility through a wholly empirical causality. But Kant apparently arrived at transcendental idealism in part by thinking of it as a revised version of the metaphysics of physical influence between substances that he derived from Crusius. Thus sensible intuition is sometimes thought of as the affection of our senses by an object not as an appearance but as a thing in itself, and transcendental idealism is thought of as having to claim (inconsistently) that we are to regard ourselves (as things in themselves) as being metaphysically influenced by things in themselves. Such a metaphysics would of course be illegitimately transcendent by the standards of the *Critique*, but Kant unfortunately appears sometimes to think that transcendental idealism is committed to it, and many of his followers down to the present day seem addicted to the doctrine that appears to be stated in the letter of those texts that express that thought, despite the patent nonsense they involve from the critical point of view. The thing in itself is then taken to be this transcendent cause affecting our sensibility as a whole, and the appearance is seen as the ensemble of representations resulting from its activity on us.

The first way of arriving at the concept of a thing in itself, which seems motivated by the basic idea of transcendental idealism – namely, its conception of our cognition as limited by conditions of sensibility – requires a thing in itself to be identical with the appearance to which it "corresponds", while the second way, which seems motivated by the transcendent metaphysics Kant was trying (apparently with imperfect success) to give up, demands that it be diverse from this thing, so as to serve as its external cause or ground. The two ways therefore involve two different (apparently mutually inconsistent) accounts of the relation between appearances and things in themselves, and two quite different constitutions, so to speak, for the regime to be established by the Copernican revolution in philosophy.

We cannot choose between the two interpretations on textual grounds, because both are clearly there in the texts. Kant even appears to

hold the obviously untenable position (which one also finds quite often among his apologists) that they come to the same thing. Moreover, both ways of talking pervade the entire *Critique*, so the problems I am raising are not the result of being persnickety about a few careless formulations. Both ways of talking are also clearly present in both editions, so we cannot resolve the issue by supposing that Kant changed his mind, shifting decisively from one to the other. I readily concede that the resulting hermeneutical situation would be best resolved if we could find a single interpretation that reconciled Kant's different statements, and it should be only with great reluctance that we should regard Kant as having committed himself to two quite incompatible doctrines. Yet with all due reluctance, we must admit in the end that this is the situation with which the texts do confront us, and there is no easy way out of it. There are even intelligible motivations for each of the two incompatible doctrines Kant endorses in the texts, so we at least have some reason for thinking that Kant's inconsistency is not a result of carelessness. The best explanation for it, I believe, is that in transcendental idealism Kant was introducing a radically new way of thinking about our cognition and its relation to its objects. Inconsistencies in his position that may seem clear enough to us now, after two centuries of attempting to come to terms with this new way of thinking, may not have been apparent to him.

However, if we are to understand transcendental idealism as an intelligible and self-consistent doctrine at all, we cannot have it both ways: we must choose. As I have already begun to indicate in the last paragraph, I think we should choose the identity interpretation and reject the causality interpretation, both because the identity interpretation is in itself a more plausible and less problematic philosophical doctrine, and because it does a better job of articulating and defending the principal claims for whose sake Kant wants to be a transcendental idealist in the first place.

Advantages of the identity interpretation

Readers of the *Critique* who take Kant's transcendental idealism to be only a minor variant on Berkeleyan idealism are probably understanding Kant through the causality interpretation. This interpretation still makes Kant's position quite distinct from Berkeley's, of course, since Berkeley takes the cause of our sensible ideas to be God and not a world of things in themselves. But both for Berkeley and for Kant (on the causality interpretation), ordinary sensible things are merely mental entities having no existence at all in themselves; to common sense, both positions seem to deprive the material world of its reality and substance. No stones are left to kick, only ideas (and even the foot with which we kick this idea is only an idea). By contrast, on the identity interpretation

Kant is saying that there is a world of things existing in themselves, and these very things are the objects of our knowledge – they are the material substances with which we interact in everyday life and which we study in the sciences. They are real, and exist in themselves independently of cognition of them. But our knowledge of them is restricted in important respects by the way we cognize them, and so there is reason to refer to these objects in a distinctive way (as 'appearances') when we have reason to call attention to these limitations (as we do when carrying out a critique of pure reason). Metaphysically, such a position is much closer to common-sense realism than it is to Berkeley's idealism. Epistemologically, however, it is still quite distinct from common-sense realism in the determinate restrictions it imposes on our cognition. Because it says that we cognize these real, material things only as they appear in space and time, which are not objects in themselves but only forms of our sensibility, Kant is also still justified in calling his doctrine a form of *idealism* – 'transcendental', 'formal', or 'critical' idealism. But it should not remind us of *Berkeley's* idealism at all.

Kant regards the objectivity of knowledge as a function not of the independent existence of objects, but of the universal validity of the concepts and laws that govern them. On the causality interpretation, this might look like a form of phenomenalism (akin to Berkeleyan idealism), that attempts to reduce real things (material objects) to patterns of sensation or tries to analyze propositions about objects logically into propositions about past, present, or hypothetical future sensations. Kant does think that the possibility of experience imposes limits on what the contents of our sensible representations can be, but his account requires that those limits be articulated not directly in terms of the sensible representations themselves, but in terms of the nature and law-governed constitution of a world of cognizable objects to which those representations provide us access. This is especially clear in the Refutation of Idealism, which claims that even the temporal order of our subjective representations makes sense only by referring them to a world of empirical objects, material substances that are *distinct* from our subjective representations. This means that for Kant, as for common-sense realism, "existence independently of the subject" is an essential part of the concept of objectivity, even if for Kant it is a derivative feature of objectivity rather than the basic idea. The identity interpretation of transcendental idealism brings this important aspect of Kant's position to the fore, while the causality interpretation tends to deny or at least obscure it.

Problems for the causality interpretation

On the causality interpretation, the existence of things in themselves easily looks like a problematic or even extravagant metaphysical claim

(as indeed it is for an idealist such as Berkeley). Or if the notion of the thing in itself is admitted at all, one may wonder (as Schopenhauer did) whether it makes sense to suppose that there is more than one such thing. Defenders of the causality interpretation must then either entertain the possibility that transcendental idealism commits us to some sort of metaphysical monism (reminiscent of Spinoza) or else find some way of justifying the thesis of transcendent metaphysics that there are a plurality of things in themselves (reminiscent of Leibniz). How this choice can be made without falling into what Kant calls "enthusiasm" (*Schwärmerei*) is something partisans of the causality interpretation have never been able to tell us.

For the identity interpretation, however, "things" are always first identified and individuated as appearances – as knowable, real, material objects having an existence independently of our subjective representations. Kant's refutation of (problematic or dogmatic) idealism (especially in the second edition of the *Critique*) even establishes that such an object must be distinct from any of my representations (KrV B xxxix). Each of these things also has an existence in itself because it can be considered apart from the relation to our cognitive faculties that makes it an appearance. It is even contained in the very concept of an appearance that it also has an existence in itself – or as Kant puts it in the second edition's Preface, if we denied ourselves the capacity to think of things as they are in themselves "there would follow the absurd proposition that there is an appearance without anything that appears" (KrV B xxvii).

F. H. Jacobi famously accused Kantian transcendental idealism of inconsistency because it maintains that the categories (in particular, the category of causality) are applicable only to appearances, yet with the doctrine of the thing in itself Kant applies this category not to appearances but to the transcendent objects that cause them.[7] It should be clear that this criticism (cogent or not) can apply to transcendental idealism only on the causality interpretation, and cannot even be articulated on the identity interpretation, which does not say that things in themselves cause appearances. It is sometimes suggested that Kant has an analogous problem on the identity interpretation because he must apply the concept of "identity" to the relation of appearances to things in themselves. But this objection fails to see that there is a fundamental difference between a *category* (such as causality) and what Kant calls a "concept of reflection" (such as identity) (KrV A260–262/B316–318). The former applies directly to objects; the latter indicates instead only a relation of our cognitive power to them. We call something an appearance insofar as it can be intuited by us and therefore cognized through our understanding; but we can *think* the same thing while abstracting from the relation to our faculties that makes it a possible object of cognition. We employ the concept of 'identity' not in order to connect two (different)

objects (as by a causal relation), but only to compare in reflection two ways in which we consider some object. In the case we are considering, this object is first introduced as an appearance, such as a material object in space and time. Abstracting from its relation to our sensible intuition, we then consider it differently, through pure understanding, as it is independently of our capacity to be sensibly affected by it. No violation of any critical strictures is involved in employing a concept of reflection in this way.

Finally, there is a fatal problem with the causality interpretation when it is applied to *myself*. Kant's famous doctrine that I am free as a noumenal being must be understood as saying that the being that is free is different from (and the cause of) the empirical I which is the only self of which I can have any empirical knowledge. Some defenders of the causality interpretation who are acute enough to see this, therefore rightly want to back off that interpretation in the case of the self.[8] Others, piling metaphysical monstrosity upon metaphysical monstrosity, speculate that the self must be regarded as a duplex, an entity having two "parts" – one of which, the noumenal self, is the unknowable cause of the other self. They do not explain, however, how these two distinct entities, standing in a cause–effect relation, can also be supposed to stand in a whole–part relation, or how they can constitute a single self any more than any other two distinct entities that stand to one another in the relation of cause and effect can constitute a single entity. Nor could they give any arguments at all on this point without violating the critical philosophy's strictures that deny us any cognition of transcendent reality.

Problems for transcendental idealism on the identity interpretation

In making the above philosophical arguments in favor of the identity interpretation, I do not mean to suggest that transcendental idealism on the identity interpretation is wholly free of problems. But I think the problems with the doctrine on this interpretation are simply the genuine problems (perhaps not insoluble) that go along with transcendental idealism itself, and with the other Kantian doctrines that properly motivate transcendental idealism. They are therefore problems for transcendental idealism, but not problems *for the identity interpretation*.

It might be thought, for instance, that there is a problem for the identity interpretation lying in the fact that appearances are essentially spatiotemporal, and yet are supposed to have an existence in themselves, even though space and time, as forms of sensibility, are not supposed to be anything in themselves. But this is nothing but a consequence of Kant's position that space and time, unlike the things that appear in them, are not objects – they are instead modes of epistemic access to those objects for beings like ourselves, for whom the intuitive component of cognition

consists in being sensibly affected by objects. It therefore lies at the heart of Kant's transcendental idealism that space and time do not exist in themselves but spatio-temporal objects do.

Transcendental idealism has some distinctive epistemic commitments that need to be clearly articulated, and the identity interpretation makes it easier to articulate them. If the distinction between appearances and things in themselves is to be tenable, then it must be intelligible to claim not only that we can be aware of our own cognitive faculties as limited, but also that we are capable of *thinking* of objects of which we deny that we can have cognition. Hegel questioned the intelligibility of both, on the ground that all thought is already a kind of cognition, and positing a limit on our cognition is already thinking (hence knowing) what lies beyond the limit.

An even more serious challenge to Kant's transcendental idealism, in my opinion, arises from the possibility that we might recognize that our knowledge is limited, but be unable to specify the limitations precisely (as Kant thinks we can do in terms of boundaries of sensibility). Without such a precise specification, it would be impossible to give a determinate sense to the notion of an object of cognition insofar as it falls within the limits and insofar as it falls outside them. In that case, there could be no determinate meaning to a distinction between things "as they appear" and "as they are in themselves." For the distinction between considering a thing "as we can cognize it" and "as it transcends our cognition" could then not yield determinate conceptions of what might be true of the thing as it is being considered or referred to in each way.

So transcendental idealism on the identity interpretation does involve a substantial philosophical thesis, which may be highly controversial. Yet this controversial thesis is simply the central idea of transcendental idealism itself – namely, the idea that we can precisely identify the sources and limits of our cognition, and on that basis can specify not only what synthetic *a priori* cognition we can have, but also which metaphysical questions about things we can never hope to answer. If Kant is wrong about that, then transcendental philosophy itself is impossible, and we should neither expect to, nor need to, articulate transcendental idealism intelligibly. On the causality interpretation, however, we could still state transcendental idealism quite clearly even if we could not draw a clear distinction between what we can cognize and what we cannot, since we could still assert that there is a transcendent cause of the objects we experience. In other words, on the identity interpretation, but not on the causality interpretation, the intelligibility of transcendental idealism stands or falls with precisely those epistemological doctrines that motivate transcendental idealism. That is perhaps the strongest argument of all in favor of the identity interpretation and against the causality interpretation.

The causality interpretation is philosophically untenable because it misrepresents transcendental idealism as a novel and extravagant metaphysical doctrine at odds with common sense – like the theocentric idealism of Berkeley or Leibniz's poetic metaphysics of windowless monads. The identity interpretation, by contrast, helps us to recognize the real nature of transcendental idealism, by presenting it metaphysically as a form of unremarkable realism, but one conjoined with the distinctive epistemological thesis that our knowledge is subject to limits that can be specified precisely enough to distinguish two fundamental ways in which the objects of our cognition may be considered or referred to – as they fall within the limits imposed by our faculties (and therefore as they can be cognized by us) and in abstraction from their relation to our cognitive faculties (and therefore as they are "in themselves").

The same distinction turns out also to set crucial limits to our cognition in another way. For it turns out, Kant thinks, that we are driven to form certain concepts (namely, ideas of reason) whose objects, if they have any, could not be presented to us in any sensible intuition. Kant thinks it is crucially important to understand where those concepts come from and to recognize that our attempts to obtain knowledge of their objects through *a priori* arguments is forever doomed to failure. Those aims, carried out in the Transcendental Dialectic, constitute his "critique" of "pure reason" in the most proper sense.

II ideas of reason

Transcendent metaphysics

Metaphysics as Kant knew it in the Leibnizian–Wolffian tradition consisted of three sciences: rational psychology, rational cosmology, and rational theology. They claimed to possess some *a priori* knowledge about the soul, the world, and God. In particular, they claimed to be able to demonstrate three crucial claims of moral-religious import: the immateriality and natural immortality of the human soul, freedom of the will, and the existence of a supreme being. Kant regarded the three "rational" sciences as pseudo-sciences claiming knowledge of matters that lie beyond human cognitive capacities. But his attitude toward rationalist metaphysics was a complex one. He thought that the questions they were asking are inevitable ones for rational human beings to ask, and he thought the metaphysical propositions they wanted to defend represent genuine human concerns from the standpoint of morality. Moreover, he thought that even the arguments the metaphysicians use are inevitably tempting to us, even if their rational force is in the end seen to be illusory. His criticism of rational metaphysics in the Dialectic

of Pure Reason is therefore far more than merely a rejection of those arguments. It is at the same time a theory about the nature and vocation of human reason, and about why human reason (as he puts it at the beginning of the Preface to the *Critique*) "is burdened with questions which it cannot dismiss, since they are given to it as problems by the nature of reason itself, but which it also cannot answer, since they transcend every capacity of human reason" (KrV A vii).

Reason's task

Even our possession of the *concepts* of an immaterial soul, a free will, and a divine being requires some explanation, since these concepts do not come to us from experience and in fact we can never experience anything that could be adequate to them. Kant's first task, therefore, is to understand where we get such concepts – or "ideas," as he calls them (with a deliberate allusion to one of Plato's terms for his supersensible "forms"). This explanation depends on Kant's conception of our faculty of "reason" itself – on its tasks and its way of carrying them out.

In the logic of Kant's day, "reason" was regarded as the faculty for drawing inferences, and hence of completing sciences by means of syllogisms. Kant regards "reason" as the "highest" or most encompassing of human faculties, the faculty through which we direct all our faculties (including reason itself) – which is why a critique directed to exploring the limits of reason must also be carried out by reason. Reason directs activities by being the faculty of "principles" – the faculty that gives the laws and rules that ought to govern both our practical conduct and our theoretical inquiries.

In its theoretical employment, reason's task is to systematize our knowledge and thus maximize the intelligibility of the world we know. We have seen that certain kinds of intelligibility, that involved in a causally governed system of substances in space and time, is required for there to be any experience at all. But beyond this there are various kinds of contingent intelligibility that may or may not be found in the world, or may be found in it only to a certain degree. They include the taxonomic intelligibility of natural kinds and the teleological intelligibility of living organisms, as well as the greatest possible completeness in the spatial, temporal, causal, and other kinds of order that constitute empirical nature. Kant understands the theoretical task of reason as that of projecting the kinds of contingent order for which we seek and directing our inquiries into it.

Kant thinks that "ideas of reason" – *a priori* concepts that, unlike the categories, cannot be exemplified in any possible experience – are generated by reason itself in the course of projecting its theoretical ends. Ideas arise when reason attempts to think in their completeness the

unconditioned conditions for what can be given in experience only under certain conditions. For example, every change in the state of any material substance in the world comes about through a causal law from a prior state of a substance, and this in turn is caused by another state, and so on indefinitely. In thinking any regressive causal series of this kind, our reason seeks an unconditioned completeness in the series, which could be found only in some state of a substance that begins of itself and requires no further cause. This, according to Kant, provides us with the idea of a *free* cause. This is an "idea" because it is a concept generated *a priori* by reason, to which nothing given in the sensible world does (or even could) correspond. Kant thinks reason likewise generates the idea of God, or of a supremely real or supremely perfect being, when it attempts to think the unconditioned totality of all the positive properties that could be found in all things – an idea which provides the material conditions for the inner possibility of all possible things, since any possible thing is simply some limited combination of the perfections found in our idea of such a supreme being.

Kant considers ideas of reason in two respects: a negative one and a positive one. Negatively, they lead to a "dialectic" or "logic of illusion" because of the tendency of human reason to pass from their status as representations to consideration of the objects they seem to represent. The compelling grounds for us to form the ideas are taken as evidence for the existence of their objects, and the characteristics we ascribe to the ideas owing to their origin in reason are taken as cognitions of these objects. We will consider the forms taken by this dialectic in the next chapter. The remainder of this chapter will deal with the rational origin of the ideas, and also with their positive function as guides to inquiry. It is easy to overlook the latter topic, or to consider it as unimportant by comparison to Kant's criticism of the dialectic of reason. I hope to correct our tendency to do this by treating the positive function of the ideas first (admittedly, out of the order in which Kant himself discusses them in the *Critique*).

The origin of the ideas in reason

Kant organizes both his discussion of the ideas of reason and the dialectic they engender into the same tripartite division of rational psychology, cosmology, and theology that he found in Wolffian metaphysics. But he seeks the principle of this order not in the contingent history of philosophy but in the logical forms constituting reason's activity itself – specifically, in the three forms of syllogism (categorical, hypothetical, disjunctive).

The doctrine of the soul arises from attempts to think a series of categorical syllogisms back to their unconditioned presupposition – namely, the concept of something that can never be a predicate of anything else

but only a subject – and this is the 'I' or self in which all thoughts must inhere, and whose necessity as their subject provides all experience with its ultimate unity. The rational doctrine of the world proceeds regressively using the indispensable relations of condition and conditioned that we find in the world – spatial and temporal relations, relations of whole to part, the relation of cause and effect regarding events, and the relation of grounded to grounding existence with regard to things. These inferences lead to the ideas of a first event in time, an uttermost limit of the world in space, a spontaneous or uncaused cause of events, and a necessarily existent being. Finally, the rational doctrine of God arises from the thought that every possible individual thing is distinguished from every other by that precise combination of affirmative and negative properties that constitute it, which could be represented by a disjunction of all affirmative properties, and the exclusion of those that do not belong to the thing. The presupposition of this disjunctive syllogism would be the concept of a single sum-total of all positive realities, which leads to the concept of a possible being having all of them and lacking none – an *ens realissimum* (most real being) or God.

Now that we no longer take the late scholastic formal logic of Kant's day to be any sort of definitive account of the actions of reason in general, it is difficult not to dismiss Kant's organization of the Dialectic as strained and artificial. But it is important for our understanding of Kant's conception of his project that we not dismiss it too quickly, before recognizing the philosophical stakes that are involved in presenting things in this way. It was Kant's fundamental aim in the *Critique of Pure Reason* to ground metaphysics as a science. His closest model for this science was scholastic logic, which he thought had achieved the status of a closed and definitive body of knowledge by limiting its pretensions and systematizing what it had rendered indisputable through this act of self-limitation (KrV B viii–ix). Kant's aim in the Transcendental Analytic is to present and justify the fundamentals of the limited and definitive system of synthetic *a priori* cognition of nature, which is presupposed by any possible empirical science of nature. Likewise, in the Transcendental Dialectic, his purpose is to present the insoluble *problems* of reason as a closed and limited system, together with the critical solution to these problems that will lay them forever to rest (at least as attempts at theoretical cognition of reality). "In this business I have made comprehensiveness my chief aim in view, and I make bold to say that there cannot be a single metaphysical problem that has not been solved here, or at least to the solution of which the key has not been provided" (KrV A xiii). For Kant, the "architectonic" use of the table of categories and types of syllogism to generate and organize the material of both the Analytic and the Dialectic is essential to vindicating the completeness and definitiveness of metaphysics as a closed and finished science.

The regulative use of the ideas

The ideas for Kant are also to play a key role in the organization of *empirical* knowledge. Kant regards experimental science as possible only against a systematic background of knowledge, which enables us to devise the questions to which experimental results give us the answers. For to be taught by nature we must, he says, approach nature "not like a pupil, who has recited to him whatever the teacher wants to say, but like an appointed judge who compels witnesses to answer the questions he puts to them" (KrV B xiii). More generally, Kant holds that empirical science differs from random groping and the haphazard collection of facts through being organized according to a rational plan. The fundamental aims, and even the structure, of science are to be determined *a priori* by reason's goals in inquiring. And the ideas of reason are vital in setting these goals.

Ideas play a variety of "regulative" roles in inquiry. Since they are arrived at through a regressive series of syllogistic inferences from what is conditioned to its ultimate (or unconditioned) condition, the process through which they arise is already one in which inquiry seeks a completed whole of cognitions. The idea thus represents that whole, and arises through reason's aspiration to know everything within it. Thus the series of past events in time, through which arises the idea of a beginning of the world, amounts to the endless aspiration to know as many of those events as possible. The regress of causes, through which there arises the idea of a first or free cause, represents the aspiration to have an unconditionally complete explanation for any event. The idea of a complete determination of all the determinations or predicates of an individual substance, through which we obtain the idea of a supremely real being or God, sets us the infinite task of coming to know all the determinations of any empirical object with which we are acquainted (KrV A670–680/B698–708).

As well as completeness, an idea involves the aspiration to unity within our cognitions. As Kant presents it, this aspiration takes several forms. It leads us to look for the minimum number of elements or forces in nature, and to seek to reduce the powers of the mind as far as possible to a single fundamental power (KrV A645–649/B673–677). It leads us to seek a taxonomy in nature according to which species are subsumed under genera, genera under still higher families, and so on, and in which all species are further subdivided by determinate differentiae down to the properties of individuals. This leads to what Kant calls the laws of "homogeneity," "specification," and "continuity" in our empirical classification of individuals under natural kinds (KrV A651–668/B679–696). The first law says that for different species there is always sameness in kind under a higher genus, the second that what is different

among members of a kind falls under lower subspecies, and the third that there is always a systematic arrangement of kinds, both higher and lower, because they are all collectively descended from a highest genus. Kant sees these laws as closely related to familiar heuristic principles, such as the scholastic dictum (sometimes called "Ockham's razor") "Entities should not be multiplied without necessity" (KrV A652/B680) and what Kant regards as its complement, "The variety of entities is not to be diminished rashly" (KrV A656/B684).

Finally, the idea of God leads us to look at the natural world as a system of ends, maximizing the intelligibility of individual living things as organic unities, and also the systematic unity of all ends within nature (KrV A685–697/B713–725). Kant emphasizes that this use must be critical, in that the regulative use of this idea must never be taken as a theoretical proof for the existence of its object, nor should the heuristic assumption that everything in nature belongs to a system of ends be used as if it were itself an explanation – as though an abstract appeal to God's will could render natural events intelligible to us. The aim is always to prompt us to seek for further empirical connections between things in the world, so as to maximize the systematic intelligibility of our experience (KrV A697–702/B725–730). This aspect of theoretical inquiry was revisited, clarified, and reconceptualized by Kant subsequently in the *Critique of the Power of Judgment* (1790).

The goals of inquiry, as Kant presents them, are given by reason *a priori*, and are not dependent on the way the reality to be known is empirically constituted. But the method of inquiry Kant projects for reason does involve a certain commitment to the way the world is. In following the regulative principles of reason, we are to look at the world *as if* the principles of homogeneity, specificity, and continuity obtained everywhere in it, and as if it were the creation of a supremely wise God, whose ends and perfectly adapted means to them were present everywhere it is possible for them to be. But of course we can never know to what extent these heuristically adopted hypotheses actually correspond to the world as we find it. Rather, Kant thinks that we are justified in assuming (for the purposes of inquiry) that the world is maximally intelligible to reason, because this is the assumption which will best promote our discovery of whatever intelligibility may be there. However, precisely because we can never know the extent to which the world corresponds to our regulative assumptions, it is crucial that we not succumb to the temptation to treat these assumptions as though our rational grounds for making them were justifications for believing that they hold, or that they ground claims to knowledge that the world is the way we rationally assume it to be.

Much of the Transcendental Dialectic is occupied with pruning back the pretensions of theoretical reason – discrediting traditional

metaphysical proofs for the existence of God and the immortality of the soul, and insisting that human cognition is limited to the realm of experience. Because of this, it is easy to read the Dialectic as expressing Kant's support for *empiricism* in the context of early modern skirmishes between it and rationalist metaphysics. No doubt there is a lot of truth in such a reading. But it should not lead us to ignore the respects in which Kant is sympathetic with the aims of traditional metaphysics, and especially it should not obscure the extent to which Kant's entire theory of knowledge is profoundly anti-empiricist in its picture of scientific inquiry and scientific theory. Kant stands at the very opposite pole from those who would portray empirical science as simply the accidental gathering of data, which is then to be ordered inductively according to the most convenient scheme we happen to think of. On the contrary, he sees science as from the beginning a product of reason, guided by *a priori* principles both in setting its aims and in setting the guiding principles for making its observations and its systematic presentation of their results. The ideas of reason for Kant are thus far more than occasions for metaphysical error. On the contrary, properly understood and properly employed, they are indispensable to empirical science as Kant conceives of it. If we are not to misinterpret Kant's theory of science quite badly, we must not ignore the positive contribution he thinks they must make to our knowledge of the empirical world.

further reading

Frederick Beiser, *German Idealism: The Struggle Against Subjectivism*. Cambridge, MA: Harvard University Press, 2002.

J. N. Findlay, *Kant and the Transcendental Object*. Oxford: Clarendon Press, 1981.

Rae Langton, *Kantian Humility*. Oxford: Clarendon Press, 1991.

Susan Neiman, *The Unity of Reason*. New York: Oxford University Press, 1994.

notes

1 Kant holds that empirical apperception (or empirical self-awareness) involves 'self-affection' – that is, being in a state that the self causes in itself. This means that on any interpretation of transcendental idealism, there has to be some causal relation involving the action of some things on themselves. But the doctrine of self-affection is not about the causing of the empirical self by another entity. This is the causal relation involved in the causality interpretation of transcendental idealism. And it entails that appearances, as effects, must be entities distinct from things in themselves as their causes.

2 Some examples of writers who take this position are: Henry E. Allison, *Kant's Transcendental Idealism* (New Haven: Yale University Press, 1983), Robert M. Adams, "Things In Themselves," *Philosophy and Phenomenological Research* (1997); Richard Aquila, *Representational Mind* (Bloomington:

Indiana University Press, 1983), Carl Posy, "Brittanic and Kantian Objects," in B. den Ouden and M. Moen (eds.), *Essays on Kant* (New York: Peter Lang, 1987).

3 Thus among the scholars mentioned in note 2 above, Allison takes the identity interpretation and Adams the causality interpretation; in each case, the argument that Kant's two ways of talking are only two ways of saying the same thing is actually a ploy in service of excluding the rejected interpretation by attempting to massage the recalcitrant texts in the desired direction.

4 This way of escaping the problem is hinted at by Sebastian Gardner, *Kant and the Critique of Pure Reason* (London: Routledge, 1999), pp. 290–8, but he does not fill out the suggestion in the way I am about to.

5 Peter Geach, *Reference and Generality* (Ithaca: Cornell University Press, 1962) and "Identity," *Review of Metaphysics* XXI (1967–8), pp. 3–12.

6 Geach's theory of "relative identity" was soon convincingly rejected by David Wiggins, *Identity and Spatio-Temporal Continuity* (Oxford: Oxford University Press, 1967) and John Perry, "The Same F," *Philosophical Review* LXXIX, No. 2 (1970), pp. 181–200.

7 F. H. Jacobi, *David Hume on Faith, or Idealism and Realism: A Dialogue* (1787), in George diGiovanni (ed. and tr.) *Friedrich Heinrich Jacobi: The Main Philosophical Writings and the Novel* Allwill (Montreal: McGill-Queen's University Press, 1994), pp. 336–8.

8 For example, Adams, in "Things In Themselves," see note 2.

chapter 5

the transcendental dialectic

Ideas of reason generate a "dialectic" or "logic of illusion" because our faculty of reason has a tendency to treat the concepts it generates as if they provided cognitions of the objects that might be thought through them, even though sensible intuition of an object is an indispensable condition for any cognition, and ideas are such that no sensible intuition corresponding to them could ever be given in our experience. The resulting illusion, Kant thinks, is not an error of particular philosophers but lies in our faculty of reason itself, which mistakes the necessity with which it forms certain concepts in the course of regulating inquiry for the givenness of objects corresponding to those concepts. Human reason itself is therefore afflicted with a "dialectic" or logic of illusion, which taunts it with the prospect of knowing what it can never know. This is like an optical illusion, moreover, in that it does not simply disappear or cease to tempt us toward error even when it has been exposed. But reason also contains the capacity to criticize the illusion and prevent itself from succumbing to the inevitable temptation. For Kant the most essential drama of philosophy is this struggle of reason with itself, and this is why he entitles its fundamental work 'The Critique of Pure Reason' – in other words, it is reason's own criticism, which triumphs over the illusions of which reason itself is the author.

There is something strange in the thought that there might be a "logic of illusion". For a "logic" involves the rules that govern a faculty of thinking and tell us how we ought to think if we are to employ this faculty properly. How could it be possible, then, for these very rules to lead us astray, into illusion, even perhaps into error? But we know from experience that it is not impossible for properly functioning faculties, for example, visual faculties, to be subject to illusions. People who seem to see a body of water at a distance in a hot desert, or to whom one of two lines of equal length looks longer in a Müller-Lyer diagram, are not suffering from any defect in their visual faculties, nor are they misusing their faculties in any way. It is even entirely possible that only someone

whose faculties are defective could be exempt from the illusions: perhaps someone who did not see water in the desert would also be unable to distinguish between real land and real water at a distance, and anyone who did not see the lines in the Müller-Lyer diagram as unequal would also be unable correctly to perceive perspectival drawings. Likewise, it is perfectly intelligible for Kant to argue that our faculty of reason, when it functions properly, makes us subject to certain conceptual illusions or sophistical lines of reasoning, and that someone who had no susceptibility to these logical illusions would not be employing the regulative principles of reason as they should. Of course, Kant also ascribes to reason the capacity to understand and criticize the illusions to which it is subject. Reason is the highest faculty precisely because it is the paradigmatic, or even the sole, *critical* faculty. Reason is capable of, and charged with, the discipline and correction of all our faculties. Reason is what prevents us from being deceived by optical illusions, by our feelings and desires, by contingent logical errors of the understanding, by the corrupting deceptions practiced on us not only by others but even more often by ourselves, and even by the necessary illusions to which reason itself is subject.

We saw in the last chapter how Kant thinks reason arrives at its ideas, and how he thinks their regulative use is indispensable to theoretical inquiry. The ideas are arrived at by rational argument – by regressive synthesis based on one of the three syllogistic forms, leading to a determinate concept of an unconditioned regarding a series of conditions. In Kant's view it is essential to reason's task of providing systematic unity to our cognitions that we unify such regressive series of them around an appropriate idea as a "*focus imaginarius*" directing both our inquiry and the organization of its results. It is easy to see how the argument for the rational indispensability of such an idea might begin to look to us like an argument for the existence of its object. This is especially true to the extent that each idea puts a unique end to a regressive series, thus apparently representing a unique object. This uniqueness can easily come to feel to us like the singularity characteristic of an intuitive cognition. Hence even though all our intuition is sensible, and thus we can have no intuition of any object of an idea generated solely by reason, it can easily seem to us that the uniqueness or singularity of the object of an idea provides us with a more than adequate substitute for the sensible intuition we need to cognize an object.

In these ways, anyone who arrives at ideas of reason and employs them in inquiry as reason itself requires, may easily fall under the illusion that the nature and function of these ideas provides us *a priori* with a cognition of their objects, assuring us of the existence of those objects. In fact, anyone who is not susceptible to this illusion would not be thinking as the highest unity of reason itself bids us to think. This is

what makes the ideas the focus of a necessary dialectic, or logic of illusion. The task of reason regarding this dialectic is not simply to avoid it – since that would be at the same time to avoid a path of thinking that is required by reason. Instead, the right path is to arrive at the ideas, to understand why they are inevitable concepts for reason to think and inevitable sources of dialectical illusion, and then to use this understanding to protect yourself from the errors to which the illusion exposes you.

I the *a priori* doctrine of the soul

Traditional metaphysics attempts to demonstrate about the soul that it is a substance, that it is the same substance over time, and that it is simple rather than composite. From the fact that a substance is a basic part of nature, and from the fact that only composites can be naturally destroyed (by being taken apart), the rational doctrine of the soul attempts to prove that the soul is immortal in the sense that it cannot be destroyed by any natural process (such as the death of the body whose soul it is). Kant discusses these arguments in a section of the Dialectic called the "Paralogisms of Pure Reason." The term 'paralogism' (in scholastic logic) refers to a syllogism that is formally invalid. Kant thus aims to show that the inferences through which rational psychology attempts to demonstrate its conclusions about the soul are fallacious. His diagnosis of the fallacy is that metaphysics treats certain features of the I, features belonging formally to its role in providing unity to experience, as if they were properties of a thing that is given in intuition as an object of cognition. He contends, however, that the I is not given to us in any such way. 'I' is merely a placeholder for whatever it is that makes our experience possible by performing the activities of understanding – synthesizing representations into the transcendental unity of apperception that makes experience possible. Through our awareness of our activity in constituting the possibility of experience, however, no inner self-object is given. The formal properties of the activity are understood wrongly if they are treated as if they were like the determinations that are predicable of a thing given to our cognition through sensible intuition.

The basic idea of Kant's critique of rational psychology can be grasped if we see how it works on the syllogism that supposedly proves the substantiality of the soul:

> *[Major premise:]* That the representation of which is the **absolute subject** of our judgments, and hence cannot be used as the determination of another thing, is **substance**.

> *[Minor premise:]* I, as a thinking being, am the **absolute subject** of all my possible judgments, and this representation of Myself cannot be used as the predicate of any other thing.
> *[Conclusion:]* Thus I, as a thinking being (soul), am **substance**. (KrV A348, cf. B410–411).

The major premise provides an analysis of the metaphysical concept (substance) that is to be predicated of the soul. The minor premise states about the thinking I that it fits the formula found in the analysis. The conclusion is that the metaphysical concept may be predicated of the soul (the I as a thinking thing).

The argument relies on the idea that a substance is what is *basic* – it is that in relation to which everything else changes, and on which all other things (using that term in the broadest sense) depend. But this is the role that the thinking I plays in experience: all our representations are present in and for the I; they change passing from one to the other, and each of them depends on its perception by the I, while the I persists through all of them and is their substratum. Anything that comes in any way into experience does so through its relation to the I. Therefore, the soul is to be regarded as a substance.

Where is the fallacy supposed to be? The answer to this question is not as easy as a reader of the *Critique* might wish. The official account is, at one level, clear enough. The major premise of each syllogism analyzes a metaphysical concept (substance) as a "pure category," without considering how or whether the analysis might apply to objects that can be given in our intuition (KrV A348–349). The minor premise states formal features of the thinking I as a condition of experience. But the conclusion predicates the property of the soul as if it were a property predicable of an object on the basis of information given through intuition of the object. Formally, Kant holds that the syllogism suffers from a fallacy of ambiguity (or *sophisma figurae dictionis*) – the premises think of the soul through pure categories of understanding only, while the conclusion treats the soul as if it were an object given in intuition (KrV A402).

To this the rational psychologist might naturally reply that there is no need to interpret the syllogism as fallacious. What difference does it make whether we regard the soul as given to us in intuition or in pure thought? The conclusion follows in any case. Even if this is supposed to matter, we need only regard 'the soul' spoken of in the conclusion also as an object of pure thinking and the argument is still entirely valid. Kant is sympathetic to this last response up to a point, in that he is willing to admit that the syllogism is valid, and its conclusion true, as long as the conclusion is not misinterpreted:

> One can quite well allow the proposition **The soul is substance** to be valid, if only one admits that this concept of ours leads no further, that it cannot

> teach us any of the usual conclusions of the rationalistic doctrine of the soul, such as, e.g. the everlasting duration of the soul through all alterations, even a human being's death, thus that it signifies a substance only in the idea but not in reality. (KrV A350–351)

Here Kant distinguishes two possible concepts of substance: one ("in reality") permitting inferences such as that a substance is not naturally destructible, the other ("in the idea") permitting no such inference. He will grant the rational psychologist that the soul is a substance "in the idea" but not "in reality." On one reading, then, he concedes that the syllogism itself is formally valid (and hence not a "paralogism" at all); but on that reading, he contends, an invalid inference occurs when the rational psychologist attempts to use the conclusion that the soul is a substance "in the idea" to infer its everlasting duration. So Kant's critique of rational psychology really seems to take the form of a dilemma: *Either* the syllogism demonstrating substantiality of the soul involves an invalid inference, *or* there is an invalid inference to the conclusions the rational psychologist wants to draw subsequently, for instance that if the soul is a substance then it is of everlasting duration. In either case, the conclusion the rational psychologist really covets (the immortality of the soul) can be reached only through a piece of fallacious reasoning.

To see what is going on here we must recall Kant's own conclusions about substances in nature as they were presented in the First Analogy. There it was argued that the determination of time requires that the duration of time is determinable only if all changes in time are alterations in the determinations of a substratum that persists and does not begin, cease, or alter its quantity through any of them. This requires that the category of substance, as a pure concept of the understanding, should be applicable to all objects of possible experience. But the concept to which this applies is one that has been *schematized* – that is, restricted to objects given in sensible intuition and interpreted as the persistence of the real in time (KrV A144/B183). It is only this that permits us to infer from the fact that something is a substance in the sense that it is a substratum of change to the conclusion that it is a substance in the sense of something persisting and not naturally destructible. When the soul is considered as a substance, however, these restrictions do not apply. The sense in which the soul is a substance involves no sensible intuition of an object, but only a formal condition for the unity of experience. From these formal properties of the *activity* of whatever it is that provides the formal unity to experience ("the I or he or it (the thing) that thinks" – KrV A346/B404) we learn nothing at all about the *real constitution* of this "thing." We do not know that it is something persisting through time, or a thing that is simple rather than composite, or even whether

the agent of our thoughts is the same thing from one moment to another or a series of different things.

Kant's position on the soul is far more skeptical than any view that has a widespread following today. He is neither a materialist nor a dualist. He does not think we can ever know whether the soul is material (whether thought or consciousness are bodily functions). Even if they are, Kant is sure that consciousness is sufficiently distinct from any corporeal process comprehensible by our physics that it will always be "impossible to explain how I am constituted as a merely thinking subject on the basis of *materialism*" (KrV B420). If the soul is an immaterial thing, Kant finds that there are three theories of the relation between it and the body: "physical influence" (a mutual natural causal relationship, as maintained by Crusius and Knutzen), "pre-established harmony" (as held by Leibniz), and "supernatural assistance" (the occasionalism of Malebranche and others). Kant roundly rejects all three, claiming they are all equally unprovable and even of doubtful intelligibility. In sum, Kant holds that the nature of the soul and its relation to the body are matters of transcendent metaphysics that lie entirely beyond the bounds of what we can ever know.

II the antithetic of pure reason

The part of the Dialectic that deals with rational cosmology threatens us not only with fallacious arguments, but also with outright contradictions. The Antinomies were the first part of the Dialectic to interest Kant,[1] and features peculiar to them tend to influence Kant's presentation of the other parts of the Dialectic as well. For instance, the thesis that ideas of reason arise from a regressive synthesis from conditioned to conditions applies far more obviously to the cosmological ideas that occasion the Antinomies than to the idea of the soul (in the Paralogisms) or the idea of God (in the Ideal of Pure Reason). Kant also uses his solution to the Antinomies as a kind of indirect proof for transcendental idealism – which he takes to be indispensable to resolving the contradictions with which we are threatened.

The cosmological ideas arise from the fact that in the world there are relations of dependency, in which one part of the world is *conditioned* by another. The fundamental idea here is that of a *world* or world-whole, which is internally complete in itself regarding the dependency-relations that hold between its parts. Each event in time is dependent on the one that came before, each part of the world in space on a part of the world that encloses it, each composite thing on its parts, each happening on its cause, each contingent being on other beings. These dependencies, moreover, give rise to a *series* of conditioned–condition relations – the

series of events going back in time, the series of enclosing parts of the world in space, the series of parts of parts of composites, the series of causes, the series of dependent beings. Kant holds that these series are generated by transcendental conditions of the possibility of experience, which require a condition for each existence that is conditioned in any of these ways. Regarding each of these series, the question arises: Does the series of conditioned conditions go on to infinity, or does it terminate in a first member of the series that is utterly unlike the other members in needing no further condition? The last sort of answer generates the cosmological ideas of a first event in time (a beginning of the world in past time) and a limit to the world in space, of a simple substance (or atom), of a first (or transcendentally free) cause, and of a necessary being (which exists by its own nature).

Each kind of answer gives us a different interpretation of the idea of a world-whole. But each pair of answers gives us two incompatible interpretations, between which we apparently have to choose. And whichever way we respond to each of the cosmological questions, the answer we give seems unsatisfactory. If we say that the regressive series of conditions goes on infinitely, then we seem to be saying that at whatever point we consider it, it is bound to be still incomplete, in which case the conditioned existence has not been supplied with what is sufficient for it to exist. On the other hand, if we say that the series comes to an end in an object corresponding to one of the cosmological ideas, then we seem to be committed to the existence of a being that violates a necessary law of experience – the law requiring each existence of that kind to be conditioned in the way that generates the series. The unsatisfactoriness of each alternative can be represented by an argument for and against the existence of an object corresponding to each cosmological idea. This threatens us with a set of contradictions: There *must* be, yet there also *cannot* be, a first event in time, a largest quantity of the world in space, a simple substance, a first or free cause, a necessary being.

The four antinomies (or five, since the first antinomy has both a temporal and a spatial part) may accordingly be summarized as follows:

> *Definitions:* x *conditions* y = df. y so depends on x that had x not been, y could not have been.
> x *R-conditions* y = df. There is an irreflexive and transitive relation R such that for all x and for all y, if xRy, then x conditions y in virtue of the fact that xRy.

In this case, we can also call x the 'R-condition' of y and say that y is 'R-conditioned' by x.

Now suppose there are entities called 'φs' given in our experience, that the *a priori* laws of experience are such that every φ is R-conditioned,

and that the *a priori* laws of experience are also such that we cannot encounter in experience any R-condition of a φ that is not also a φ. The *thesis* of each antinomy then asserts that

> Something that is not R-conditioned must exist as the first member of the R-conditions of any given φ.

The *antithesis* of each antinomy asserts that:

> All the R-conditions of any given φ are themselves φs, hence R-conditioned by further φs to infinity.

We may represent the four (or five) antinomies using the following scheme of values for φ and R:

Antinomy	*φ*	*R*
First (time)	event	precedes
First (space)	spatial region	properly encloses
Second	composite body	is a proper part of
Third	event	causes (according to a law)
Fourth	contingent being	grounds the existence of

Let us take the temporal part of the first antinomy as an illustration. Here the thesis says that there must be a beginning (a first event) of the world in time, and the antithesis says that every event in the world is preceded by another event back to infinity (KrV A426–427/B454–455). The argument for the thesis is that the *past* is that series of world-events which has already been completed, and an *infinite* series is one that cannot be completed; therefore, there is a contradiction in the idea of an infinite past, and hence the series of past events must be finite, or there must have been a first event to begin it (KrV A426–428/B454–456). The argument for the antithesis is that if there were a first event, then it must have been preceded by an empty time, but (on the basis of the argument for the First Analogy), in an empty time nothing can arise. Hence a first event is impossible, and so the series of events in the world must go back to infinity. Analogous arguments supposedly can be given both for and against the existence of a largest region of world in space, a simple substance, a first (or free) cause, and a necessary being.

The pull of both sides of the antinomies

There is some reason to doubt that the arguments on either side of the antinomies should convince us of anything. Regarding the general argument against the thesis of each antinomy, why must we suppose that the "conditions" relation is transitive? Perhaps each effect is produced by its

cause, and it is an altogether separate question whether or what causes that cause requires. Perhaps every temporal state is preceded by one that immediately precedes it, but why should this state of the world, and even all the states that preceded it, be *conditions* for the existence of the present one? Thus whether every φ (e.g. every event in time) is conditioned by a preceding one might be a contingent, empirical question, not something that can be settled by an *a priori* law of experience. Thus there might, for instance, just be a first event in the world (a "Big Bang") without any preceding event, or a simple substance (an atom) that (just as a matter of brute fact about the empirical laws of physics) cannot be further divided. The argument for the antithesis of the antinomies thus may not seem very compelling.

There is also something suspicious about the general argument against the antithesis, and thus for the thesis, of each antinomy. If a conditioned existence requires an infinite series of conditions, why should we see any problem with this, or worry that it threatens us with an insufficiency of conditions? After all, the actual existence of the conditioned object is pretty clear evidence that all its conditions have been fulfilled, whether they are finite or infinite in number. For example, a series of events in time can be *infinite* in any of three ways: by having a beginning and no end, or an end and no beginning, or neither an end nor a beginning. The series of future events might be infinite in the first way, the series of past events might be infinite in the second way, and the series of world-events as a whole might be infinite in this third way. This might make the thesis of the antinomies seem equally without support.

But the deeper worry may be that without an unconditioned condition, either residing mysteriously in the entire infinite series as a whole or else concentrated in some exceptional first member of it, we have not yet specified the *kind* of condition that could truly *satisfy the conditions required* for the existence of the conditioned thing. Thus a series of conditioned conditions, even an infinite one, still does not yield any thing that truly satisfies the conditions for the conditioned thing, which would mean satisfying them *unconditionally*. Thus there is a philosophical inclination, having a profound grip on us, that future events *depend on* (are conditioned by) past events in a way that past events are not conditioned by future ones. This sense of temporal asymmetry makes it mind-boggling to think of a *past* that had no *beginning* in a way that it is *not* mind-boggling to think of a *future* that will have no end. This, I think, is the nagging worry that convinced Kant that the thesis of the antinomies cannot be dismissed in the way suggested in the preceding paragraphs. The very same intuition of dependency (or conditionedness) also exercises a pull in favor of the antithesis. For if each event *depends* on (is conditioned by) there having been a preceding

event, then the very idea of a first event (or beginning) to the world may also seem inconceivable.

For this reason, Kant thinks, the antinomies leave us perplexed and dissatisfied whichever solution to them we may decide to favor. As he puts it: when we try to form a concept of these cosmological series, the thesis seems to present us with a concept that is *too small*, while the antithesis presents us with one that is *too large* (KrV A485–490/B513–518). (The exception is the fourth antinomy, where Kant thinks the idea of a necessary being is too large for our concept, while an endless series of conditioned beings is too small.) Kant does not expect that we can ever rid ourselves entirely of the sense of puzzlement and dissatisfaction occasioned by these abysmal problems. At most he hopes to resolve the issue before the bar of reason so that we can at least be freed from *error*, prevented from making judgments on one side or the other whose rational grounds are illusory rather than genuine.

Resolving the antinomies

Kant's solution to the antinomies depends on drawing a distinction between things of nature as appearances and a realm of things in themselves. Kant divides the antinomies into two groups, depending on the kind of conditioning relation they involve. The first two he calls *mathematical* antinomies, because they involve temporal, spatial, or part–whole relations between things as they are given to intuition in space and time. The third and fourth antinomies he calls *dynamical* antinomies, because they involve relations of causal dependency between happenings or between things.

The mathematical antinomies are generated by mathematical principles that apply to things only insofar as they are given in sensible intuition. As so given, however, they constitute a regressive series of conditions that is indefinitely long – but neither finitely nor infinitely long. For each event in time, it must be conditioned by an earlier one, for each extensive portion of the world in space, it must be conditioned by a larger one, and for each part of a substance having spatial extension it must be a composite conditioned by its proper parts. But these series of conditions are never given to intuition as a whole. Kant thinks that to assume they must exist either as infinite wholes or as finite wholes is to assume that they are not merely appearances but things in themselves whose determinations must exist independently of the manner in which they can be given to our intuition. But if transcendental idealism is true, this assumption is false. It follows that both the thesis and the antithesis of the mathematical antinomies are *false*. The theses are false, because the principles of possible experience make it impossible for objects corresponding to the cosmological ideas of a first event, a largest extent of

the world, or a simple substance ever to be given to intuition. The antithesis is false because there is *no fact of the matter* about the age of the world in time, its extent in space, or about whether the divisibility of composites given in experience is finite or infinite. Consequently, there can be no fact that these are infinite. The arguments for both sides of these antinomies rest on a fallacy of ambiguity similar to the one found in the paralogisms. They draw on principles that apply to conditioned existences considered as appearances given to our intuition, but they try to reach conclusions that would have to be true of these things only if they were considered as existing in themselves apart from the way they are given (KrV A517–527/B545–556).

Regarding the dynamical antinomies, Kant's solution again depends on the distinction between things as appearances and things considered as existing in themselves. But this time, he concludes not that both sides are false but that both the thesis and antithesis are (or might be) true. The thesis is false when it is applied to appearances. For no event uncaused by another and no being whose existence is independent of other beings can ever be given in appearance. But if we consider the cosmological ideas of a first or free cause and of a necessary being as referring to things in themselves (that cannot be given in experience), then there is no contradiction in supposing the existence of such things. But since they cannot be given in intuition, we could have no cognition of them and so their existence must forever remain an unsettled question, at least from the standpoint of theoretical reason (KrV A532–537/B560–565, A559–565/B587–593). Again, the arguments for both sides depend on a fallacy of ambiguity in failing to distinguish the supposed objects of the cosmological ideas as appearances and as things in themselves.

Doubts about Kant's solution

There is good reason to be skeptical of Kant's solution to the antinomies, and especially of his thesis that transcendental idealism is needed to solve them (which thesis he also presents at KrV B xvii–xxii as a kind of "indirect proof" of transcendental idealism itself). Kant's solution depends on the claim that both sides of the antinomies err in supposing that if the conditioned is given, then the totality of its conditions, hence the unconditioned, must also be given. He seems to concede that if the totality of conditions is given, then that totality would have to be either finite or infinite in extent – thus leading to an equally valid argument on each side, and thus to an irresolvable opposition between equally demonstrable contradictories. Kant's way out is to deny that the conditions (and the world, regarded as the series of conditions) can be given as a totality. This would be plausible if the claim is only that under the laws of experience established in the Transcendental Analytic, we can

have no direct experience either of a first event in time or of an endless past series of events, of an indivisible part of a composite or of its infinite division, and so on. But the natural sense of 'given' in this context is not 'directly experienceable' but rather 'existent' or 'actual' in the sense of the Postulates of Empirical Thought – namely, that something exists or is actual if it is *connected to* some intuition by (either transcendental or empirical) laws of experience (KrV A217/B266, cf. A376). This postulate of actuality is needed if Kant is to admit the actual existence of corpuscles too small to be visible or tangible by us, or celestial objects too distant ever to be visited by us, or even of most of the past, which we cannot now actually perceive or even directly remember, but must infer from its connection with directly perceivable evidence (archives, fossils, written memoirs, and so forth). But if 'given' means 'existent' in that sense, then surely 'the world' (the various series of conditions of any given conditioned) is also 'given'. The only question is whether 'the world' is really an 'object' at all – that is, whether the category of 'totality' (a pure concept of the understanding, hence a necessary concept of an object in general) is applicable to 'the world'. If it is, then it would seem that there is necessarily a world-whole (the unconditioned totality of the series of conditions) and then it is either finite or infinite. So the arguments of the antinomies threaten us with the conclusion that it must be both (hence with a contradiction).

Kant's way of avoiding the contradiction, then, comes down to the claim that the category of totality cannot be legitimately applied to 'the world' (to the various series of conditions that generate the antinomies). But it is not clear how he can avoid applying the category of *totality* to the series, any more than he could avoid applying the categories of *unity* or *plurality* to it. For surely each series is *one* series that has *many* members – and if so, why is it not a *whole* series – whose magnitude, therefore, must be either finite or infinite? It is also unclear how transcendental idealism is supposed to help out here. For why should the category of totality be less applicable to appearances than it is to things in themselves? It might be thought less applicable to appearances if we are using the notion of 'given' – as applied to appearances – to mean 'directly presentable in present (or future) experience'. But we have seen that Kant cannot consistently apply the notion of the 'given' in this restrictive way to the world of appearance as long as he wants to count imperceptible corpuscles, or distant bodies, or even the prehistoric past as belonging to the world of appearance.

The one device still left open to Kant by his official doctrines is to distinguish between two sorts of 'laws' by means of which a putative 'given' might be connected with actual perception. One sort includes both the transcendental laws spelled out in the Principles chapter of the Analytic and the empirical laws grounded on them. The other sort

includes principles not of the understanding but of reason – in particular, the principle that if the conditioned is given, then the whole (unconditioned) series of its conditions is given. This principle, as Kant rightly points out, is synthetic: "for the conditioned is analytically related to some condition, but not to the unconditioned" (KrV A308/B365). Kant's official position is that such synthetic principles of reason are only regulative and not constitutive – they instruct us how to inquire, and what assumptions to use as the basis of our inquiries, but they do not guarantee the truth of these assumptions or guarantee that the world in its real constitution corresponds to them. This distinction would permit Kant to say that the totality of each series of conditions is not 'given' relative to constitutive principles, but only assumed by regulative principles, and that this blocks the inference that the whole series of conditions must be an actually given finite or infinite whole.

Yet one of the aims of the Dialectic is to *establish* that principles of reason are merely regulative, not constitutive. Perhaps we should see the Antinomies as Kant's indirect proof of *this* claim, if its acceptance is our only way of avoiding the contradictions. On this showing, however, the role of transcendental idealism in resolving the Antinomies would seem to have vanished entirely. For if principles of reason are regulative, not constitutive, it would seem that they must be equally so when applied to appearances and when applied to things in themselves. In other words, Kant has given us no reason to think that the antinomies would be any more irresolvable if we take the world-whole to exist in itself than if we take it to consist of appearances. Thus even if Kant can successfully resolve the antinomies in something like the manner he intends, he does not seem to be correct in holding that transcendental idealism is needed to do it.

The problem of freedom

The antinomies have special interest for Kant insofar as the third antinomy in particular relates to the problem of freedom of the will, which he regards as profoundly important for the possibility of practical (or moral) reason. Kant returned repeatedly to this topic, not only adding two extraordinary sections to the first *Critique* in order to deal with it, but also devoting to it the Third Section of the *Groundwork for the Metaphysics of Morals* (1785) and large portions of the *Critique of Practical Reason* (1788), as well as revisiting it in the First Book of *Religion Within the Boundaries of Mere Reason* (1793–4). The following account is an attempt to put together what seems least unsatisfactory out of this long and restless search.[2]

Kant holds that the validity of the moral law depends on our having "practical freedom" – the capacity to act on principles we give ourselves

through reason, and to resist the pull of the desires arising from our natural needs as living beings. We are not only committed to regarding ourselves as practically free by our vocation as moral agents, but we need to think of ourselves as free even in order to ascribe our theoretical judgments to ourselves (G 4:447–448). Thus we would need to think of ourselves as free even in order to represent ourselves as judging *for reasons* that we lack freedom – a point which makes the denial of freedom self-refuting, no matter how good the arguments for it might otherwise be.

There is a problem, however, even in trying to think of ourselves as practically free without falling into a theoretical self-contradiction. For all our actions, as events in the world of appearance, fall under laws of natural causality, and are thus causally determined by natural events preceding them in time. Yet Kant sees no way in which we can be practically free unless we are able to begin a series of events in the natural world independently of any natural causes that might influence us. Consequently, he holds that we cannot regard moral laws as valid for us – we cannot regard ourselves as morally responsible beings, or even as rational theoretical judges – unless we ascribe to ourselves the capacity to be the kind of cause we conceive under the cosmological idea of a first or free cause – the very idea that is at issue in the third antinomy. But it is not clear how we can avoid an outright self-contradiction if we apply that idea to ourselves while also acknowledging that our actions are natural events causally determined by natural laws.

Kant's solution to this problem is once again to appeal to transcendental idealism's distinction between appearances and things in themselves. Determination by natural causality applies to our actions as parts of the world of appearance, but Kant holds that it is consistent with this to regard ourselves as free when we are considered as things in themselves. Since space and even time are features of things only as appearances, our actions as events in time may fall under causal regularities governing such events, yet at the same time they may fall under a intelligible causality proceeding from a timeless choice we make as members of a noumenal world.

It is important not to misunderstand this solution by mistaking its purpose or status. Kant does not think that we can ever prove theoretically that we are free, or achieve any cognition of our free actions. Since he thinks knowledge of what goes on in a noumenal or intelligible world of things in themselves is entirely impossible for us, it would be self-contradictory for him to claim that we *know* that we are free agents in the intelligible world, or indeed to make any positive claims about how such a free causality might operate. His legitimate aim can be only to show that there is nothing self-contradictory in regarding our actions as events falling within the causal mechanism of nature and also asserting that they are effects of the free causality of our reason. All he needs in

order to do this is to establish that there is no self-contradiction in supposing that we exercise free causality as noumenal beings. Once he has established the self-consistency of asserting we are free while also viewing our actions as events in nature, he can (and indeed *must*) disavow any positive account of how freedom and natural causality actually relate to one another. It is sadly true, however, that Kant seems to have thought it appropriate that in thinking of ourselves as free, we should also think of ourselves as members of an invisible world (a kingdom of God or realm of grace) hovering somewhere beyond the realm of nature. This crotchet leads him at times to attribute a sort of positive reality to the theory of free action as noumenal causality. The point to insist on, however, is that his actual doctrines do not require this indulgence of metaphysical bad taste – indeed, they do not even permit it. These doctrines allow us – indeed, constrain us – to say that we can without inconsistency regard ourselves as free and also as parts of the natural world. Beyond that, they require us to be austere metaphysical skeptics about what (or where) freedom of the will is or how it is possible.

Timeless agency, historicity, and empirically free action

It follows that Kant also can (and even must) reject any inference from the idea of free noumenal causality to conclusions about how our freedom is to be understood empirically. It is easy for Kant's critics to caricature his conception of moral agency by drawing such inferences on his behalf – as by claiming that Kantian moral agents must feel alienated from their natural existence, think of their actions as occurring outside time, and hence be unable to think of themselves as historical beings, and so on. Criticisms of this sort were mounted against Kant in his own day, and they still appeal to many at the present time. But they are all utterly worthless, since they are based on a *non sequitur*, for the reasons just given. Nothing whatever about moral psychology or our experience of the moral life, or about empirical human nature, can be legitimately inferred from Kant's critical solution to the abstract, metaphysical question of free will. That solution involves nothing beyond a proof that freedom and natural causality are logically consistent. Consequent upon Kantian principles, it cannot be understood as providing any positive account of what our freedom consists in. Perhaps merely the total absence of such an account is what leaves us in some sense deeply self-alienated. If so, then that is our condition, and we should simply face up to it. There are, after all, far worse things than being self-alienated. One of them is believing consoling lies about ourselves in order to avoid *feeling* self-alienated.

We will see in the next chapter that Kant has a definite theory of human history, based in part on rational (regulative) principles and in

part on empirical considerations. This theory consistently assumes that human beings are both natural beings and also free beings. No positive role is played in it by the picture of our freedom as a capacity belonging only to a timeless, noumenal self. In fact, it would have been inconsistent with critical principles for Kant to employ the idea of noumenal causality in any such way.

In the Canon of Pure Reason, Kant maintains that 'practical freedom' – our capacity not to be immediately necessitated by sensuous impulses, but to choose how and whether to act as they give us incentives to act – can be proved empirically, because human beings show themselves to have this capacity, whereas brutes do not (KrV A802/B830) – though he thinks this empirical point still does not remove the metaphysical grounds for doubting freedom of the will that are the subject of discussion in the Third Antinomy and its resolution (KrV A803/B831). But Kant clearly understands our empirical capacities to plan for the future, to select between alternative means to our ends, to devise our own conceptions of happiness, and to be motivated by moral considerations, to be empirical *signs* (if not empirical *proofs*) of our freedom. If there are problems reconciling freedom with our knowledge of human beings at the empirical level, then on Kantian principles these problems would also have to be solved empirically – though it must be admitted that Kant himself did not deal very much or very profoundly with these questions, perhaps because he thought the issue of free will was really only one of transcendent metaphysics.

Throughout his historical and anthropological writings, Kant does not agree with those who interpret him as holding that from a "scientific" or "observer's" or "third person" perspective, human beings must be understood as will-less automata, our actions determined mechanically like the motions of billiard balls careening about on a green felt table, while from a "moral" or "agent's" or "first person" perspective, we must understand ourselves as free.[3] He never suggests that we should (or even could) reconcile the obvious contradictions between freedom and fatalism as empirical views of human action simply by assigning the contradictory claims to different "standpoints." (One might as well try to resolve Zeno's paradoxes simply by saying that there is a "standpoint" from which things can move and a "standpoint" from which they cannot, or the liar paradox just by saying there is a "standpoint" from which what the Cretan says is true and also a "standpoint" from which it is false.) Such a "solution" to the free will problem also naturally invites the thought that the "first person" perspective is nothing but a subjective illusion, while only the "scientific" perspective (denying freedom) gives us the objective truth about ourselves. Those who try this sort of solution today usually see it as an alternative to (a way of avoiding) Kant's metaphysical thought that we belong to two different realms – a

realm of sensible appearances and a realm of intelligences. The only legitimate use Kant could make of the idea of these two realms, in keeping with the strictures of critical epistemology, would be to indicate a bare logical possibility – and thus saving freedom and determinism from outright contradiction. But even a dogmatic metaphysics of noumenal freedom would at least permit us a self-consistent and intelligible way of reconciling freedom with causal necessity. Contemporary 'two standpoints' interpretations motivated by a desire to avoid metaphysics cannot do even that.

In all his writings about our empirical knowledge of human beings and their actions, however, Kant treats the "practical standpoint" as thoroughly integrated into our (second- and third-person) objective empirical observations of human beings as rational agents, whose exercise of rational capacities is empirically observable by others as well as by themselves. He seems (wrongheadedly in my view, perhaps even inconsistently) to have thought that the metaphysical defensibility of doing this is tied to our being able to regard ourselves as members of a supernatural order of things, as though regarding ourselves honestly and soberly as part of nature were incompatible with finding moral value in ourselves. Today this position looks unattractively like the superstitious idea of those who think that there is something morally pernicious about believing the theory of evolution regarding human origins. But Kant never so lost touch with good sense as to declare that we can regard ourselves in such a way only from a first person point of view, or to think that this bizarre declaration would constitute a solution to the free will problem.

III god as the ideal of pure reason

The third chapter of the Transcendental Dialectic deals with the metaphysical pseudo-science of rational theology. This chapter of the *Critique* is famous for its criticisms of the traditional scholastic-rationalist proofs for the existence of God, especially its criticism of the "ontological proof" (a name for this argument of which Kant was the inventor). But it is important to recognize the positive as well as the negative side of the Ideal of Pure Reason as far as traditional rational theology is concerned.

Kant is famous for what Moses Mendelssohn called his "world-crushing" (*weltzermalmend*) criticisms of the traditional metaphysical proofs for God's existence. Kant has often taken the credit (or blame) for revolutionary developments in theology that occurred in the early nineteenth century, when German idealism replaced the traditional scholastic-rationalist metaphysical theology with new conceptions of divinity that

stressed divine immanence and emphasized aspects of religion allied to the aesthetic side of our nature at the expense of the metaphysical attributes of God that might be made the subjects of rational analysis and argument. In fact, however, on this point the reality of Kant's views diverges widely from their common reputation. As a rational theologian, Kant is much closer to the dry rigor of the scholastic rationalists he criticized than to the *Schwärmerei* of the Romantics and speculative gnostics who later laid claim to his philosophical legacy.

Kant's criticisms of the traditional proofs for God's existence were present already in his early treatise *The Only Possible Ground of Proof for a Demonstration of God's Existence* (1763). But as the title of this work implies, Kant thought at this time that the existence of God is demonstrable, by an argument based on the conditions for the metaphysical possibility of anything in general. He argued that the perfections of God were the indispensable material conditions for the possibility of anything, so that an assumption of the non-existence of God would entail not merely the non-existence but the absolute impossibility of anything (Ak 2:70–92). In the *Critique of Pure Reason* Kant has ceased to regard this line of reasoning as a demonstration of God's existence, but it still plays a very important role in the argument, mentioned in the previous chapter, for the claim that the idea of an *ens realissimum* arises inevitably from our attempt to think the conditions for the possibility of any individual thing as consisting in the "complete determination" of its individual concept – that is, the precise combination of perfections (or "realities") and their absences (or "negations") that go to make it up (KrV A571–583/B599–611). On the basis of this argument, Kant claims that the idea of God is the sole "ideal" of which reason is capable – that is, the sole idea of an individual thing that is completely determined through its concept alone (KrV A568/B596).

In his *Lectures on the Philosophical Doctrine of Religion*, Kant works out the idea of God in a very traditional way, as a being possessing the "ontological" predicates of omnipotence, omnipresence, immutability, and timeless eternity, as well as the "cosmological" or "anthropological" predicates of omniscience and moral perfection of will, based on our analogical predication of the perfections we find in our own understanding and volition (Ak 28:1012–1082). Kant's conception of God thus always remained very close to that of the scholastic rationalism of Leibniz and Wolff. This scholastic-rationalist metaphysical theology has often been viewed as closely allied with religious dogmatism and authoritarianism (as personified, for instance, in the character of Hume's Demea in the *Dialogues Concerning Natural Religion*, who is supposedly modeled on Samuel Clarke). If we are to understand Kant's view of the matter, however, we must entirely unlearn this set of associations. For to Kant it was very important to maintain the purity of the

concept of God as presented in the metaphysical conception of a transcendent *ens realissimum*. For only this, in his view, can guard against the "anthropomorphism" and consequent superstition and moral corruption of popular religious culture – toward which, as a representative of the Enlightenment, Kant always entertained the deepest suspicions.

The three kinds of theistic proofs

For the purposes of criticizing the traditional proofs for God's existence, Kant divides them into three kinds:

1 *Ontological proofs*, which argue for the necessary existence of a supremely perfect being solely through its concept.
2 *Cosmological proofs*, which argue for the necessary existence of a supremely perfect being from the contingent existence of a world in general.
3 *Physicotheological proofs*, which argue for the existence of a supremely perfect being from the contingent constitution of this world as we find it empirically (e.g. from the seemingly purposive arrangements we find in it).

Kant's strategy is to argue that the second and third kinds of proof cannot succeed in establishing the existence of an *ens realissimum* without relying tacitly on the first kind (the ontological proof); and that no ontological proof of God's existence is possible – thus undermining all theistic proofs by a sort of domino effect. One consequence of this strategy is that Kant in effect mounts no criticism at all of the cosmological inference from contingent to necessary existence, or from apparent purposiveness in nature to the existence of some sort of intelligent designer of the world. (In his 1763 essay, however, Kant did present such criticisms, and there is indication in the *Critique* that he did not intend to let these inferences pass unchallenged. See KrVA609–610/B637–638 and A626–628/B654–656.) A second consequence is that Kant's entire *official* critique of the theistic arguments is made to rest on his critique of the ontological argument.

It is certainly possible to doubt whether the other two kinds of theistic argument presuppose the ontological argument. In the case of the physicotheological argument, Kant's claim seems to be not that it presupposes the ontological argument, but rather that it cannot be understood as a proof of a supremely perfect being – and therefore that something like the ontological argument will be needed in any case to establish the existence of such a being (KrV A625/B653). Thus there are quite a number of issues about natural theology that Kant does not even pretend to address conclusively in the Ideal of Pure Reason. But for the rest of our discussion, we will take Kant at his word and concentrate on his famous criticism of the ontological argument for God's existence.

Is 'existence' a "real predicate"?

The ontological argument, in its simplest form, is that since God is a being of whom all perfections must be predicated, and since existence (or necessary existence) is a perfection, therefore, God must exist. If the argument is to be given its due within the metaphysical tradition to which it belongs, it has to be realized that 'most real being' and 'most perfect being' are not merely arbitrary verbal definitions, but depend on an ontology – the same ontology to which Kant subscribes in the reasoning that generates the idea of God – in which the nature of *any* entity is seen as consisting of a certain combination of realities (or perfections) and negations. Such an ontology will, as Kant fully realizes, naturally have a special place for the idea of a being having *all* realities or perfections – in fact, in such an ontology the idea of such a being will be fundamental to any conception of any thing. The ontology itself makes sense only to someone who accepts the thought that all realities or perfections can be found in the same being, and even that in their highest and most complete form, they *must* be found together in the same being. If 'existence', or especially '*necessary* existence', is one of these supreme realities or perfections, then it will not seem plausible to attach it arbitrarily to any being other than the one in which all perfections, in their highest form, are to be found. And there will be considerable intellectual appeal to the thought that there is a meeting point of the order of our concepts of things and the order of existing things (that they are concepts of), which point is to be found in that being which, having all realities or perfections, also has the perfection of necessary existence.

It is against this background that we should understand the acceptance of the ontological argument, in one form or another, by many seventeenth-century rationalists from Descartes to Leibniz. It is in the same context also, however, that we must try to understand Kant's famous critique of the ontological argument, which takes the form of denying the crucial premise that 'existence' is a 'real predicate', that is, a reality or perfection. Most of those who have agreed with Kant's critique of the argument have not understood it in this context, but have instead accepted it simply as a rejection of the entire metaphysics of realities or perfections, together with the reasonable expectations to which this scholastic-rationalist ontology might give rise. But we have seen that Kant was actually very sympathetic to that ontology, so that this way of "agreeing" with him not only amounts to a dismissal of the set of ideas that gave the argument its appeal, but is also very likely to involve a serious misunderstanding of what he had in mind.

Kant declares that "existence is not a real predicate," that is, "it is not anything that could be added to the concept of a thing" (KrV A599/B626). This of course does not mean that it is a *phony* predicate, nor does it

mean to deny that when we say truly that "X exists" we are supplying some additional information about X. Kant's claim takes for granted that the concepts of things are generally made up of "real predicates" – that is, realities or perfections as conceived in the traditional ontology. But Kant wants to draw a distinction between (1) propositions that "determine" a subject-concept by predicating some "reality" (or perfection) of it and (2) propositions that only "posit" an object corresponding to the subject concept, without predicating of it anything that could be part of the concept itself. Propositions of the form "X exists" are of the latter kind. "When we say 'God is' or 'There is a God', we attach no new predicate to the concept of God, but only posit the subject itself with all its predicates" (KrV A599/B627).

Many have thought the thesis that existence is not a real predicate is self-evident, but to someone who (like Kant) accepts the traditional ontology, it surely can be nothing of the kind. Kant has, however, remarkably little to say in defense of his thesis. It is uncontroversial enough to say that "X exists" asserts that there is some object to which the concept of X corresponds. But the point that really needs to be established is that 'is' or 'exists' is not *also* a reality or perfection, that might belong to the nature of a thing and hence be contained in its concept. This needs especially to be established if the predicate is not 'exists' but '*necessarily* exists', since it does seem that something that exists necessarily is more perfect or more real than something that exists only contingently. So it is hard to see how someone could consistently say of the *ens realissimum* that it has all perfections but lacks existence, or that its existence is only contingent, like that of less perfect things.

There is a somewhat analogous problem with emotivist metaethical theories, those that hold that 'X is good' predicates no property of X but instead only expresses the speaker's "commendation" or "approval" of X. There too it is uncontroversial that calling something 'good' normally expresses approval or commendation, but the point that needs to be established is that 'good' does not *also* refer to some real property of things. (As a matter of fact, calling something 'good' does not directly *express* approval or commendation of it at all, but rather asserts that the thing has a property such that an attitude of approval or commendation is rationally *warranted* or *justified* by its possession of this property.) The argumentative move of saying that 'existence' or 'good' are not used to predicate but rather to perform some other semantic function (of "positing" or "commending") is not one that should, in general, expect to meet with uncritical agreement. Suppose a philosopher claimed that 'heavy' is not a real predicate, and supported this by arguing that 'heavy' serves the unique semantic function of 'gravitizing' its subject, or that 'blue' is not a real predicate because it does not ascribe a property of the subject but instead 'azurates' it. Emotivists and defenders of Kant's

thesis that existence is not a real predicate need to show that "commending" and "positing" do not function in their contentions in the same (utterly unpersuasive) way that 'gravitizing' and 'azurating' do in these.

We will understand Kant's criticism of the ontological argument better if we see it as a rejection not of the metaphysical ontology of realities or perfections, but rather of the intellectualistic epistemology through which the ontological argument appropriates that metaphysical theory. For Kant, cognition requires that an object be *given* in intuition and *thought* through concepts. The modal category of *existence* applies to things by expressing the givenness of the object, the connection of it to an actual intuition (through sensation) (KrV A218/B266). It is this givenness of the object that is expressed by the 'is' or 'exists' that "posits" a subject to which the concepts of various realities may be predicated. Because intuition is a function of cognition that is distinct from conception, no concept can express this condition of cognition. Therefore, the existence of an object can never be included in its concept, but must always be added to it through an intuition in which the object of the concept is given.

For Descartes, by contrast, our idea of God is the immediate presentation in thought of a "true and immutable nature" – the nature of a supremely perfect being, in which all perfections are given to us in their indivisible unity. From our idea of such a nature we obtain cognition of the properties belonging to it by predicating of this nature whatever can be drawn from it – including, of course, the necessary existence that belongs to it along with all other perfections. In effect, Descartes regards the ideas of true and immutable natures not only as concepts but also as something like intuitions (in Kant's sense) of the objects they represent. These quasi-intuitions do not in general guarantee the actual existence of the object represented, though they do provide us with certain cognition of its predicates. Thus our idea of a true and immutable nature of a triangle enables us to know that the angles of any possible triangle are equal to two right angles, even though it does not enable us to know that any triangle actually exists (since like the nature of all created things, this nature contains only possible or contingent existence). Our idea of the true and immutable nature of God, however, contains necessary existence, and so from it God's actual existence can and even must be inferred.

Kant recognizes something like Descartes' epistemology of true and immutable natures in the form of mathematical concepts whose objects we can immediately represent through *a priori* construction in the pure intuitions of space and time. But for Kant it is essential to such concepts that their objects should be spatio-temporal, and Kant would not say of these objects that they *exist* except in the sense that instances of triangularity or the number five can be given as empirical objects *in sensation*.

> If the issue were an object of sense, I could not confuse the existence of the thing with the mere concept of the thing. For through its concept, the object would be thought only as in agreement with the universal conditions of a possible empirical cognition in general, but through its existence it would be thought as contained in the context of the entirety of experience; thus through connection with the content of the entire experience the concept of the object is not the least increased, but our thinking receives more through it, namely a possible perception. If, on the contrary, we tried to think existence through the pure category alone, then it is no wonder that we cannot assign any mark distinguishing it from mere possibility. (KrV A600–601/B628–629)

The concept of God is not one whose object can ever be given to us in any intuition, whether pure or sensible. For this reason, no genuine cognition of God is possible for us. The closest we can come to such cognition is to analyze the pure idea of a supremely perfect being to see what real predicates it contains. Kant thinks that since in this case there can be no distinction between concept and intuition, there is a temptation to treat the givenness of the object itself (expressed by "positing" this object through asserting its existence) as though this too were just one more determination (perfection or reality) belonging to its concept. This creates the dialectical illusion that we can cognize God's existence merely by analyzing the concept of God. Kant's critique of the ontological argument should be read as the attempt to expose this illusion and break its hold on us.

If this interpretation is correct, then it would be a mistake to say that Kant had succeeded in finding – or even that he pretended to find – in the ontological argument some elementary fallacy or logical error. On the contrary, Kant's critique of the ontological argument would be only as sound as the most fundamental thesis of his epistemology – that all cognition requires *both* that an object should be given in intuition and that it should be thought through concepts. This thesis is not an elementary point of logic, and there are many (even many who are entirely unconvinced by the ontological argument) who would regard it as questionable. Hence those philosophers who think there is some simple fallacy or logical blunder that vitiates the ontological argument should not cite Kant as agreeing with their thought, nor should they labor under the illusion that there is anything in Kant's writings that would lend support to it.

IV the transcendental doctrine of method

The *Critique of Pure Reason* has two main parts: The first, the "Transcendental Doctrine of Elements," is divided into the "Transcendental Aesthetic" and "Transcendental Logic" (which includes

both the Analytic and the Dialectic). The second major division, "The Transcendental Doctrine of Method," tends to be neglected by its readers (perhaps just because the book is so long and the parts already surveyed are so exhausting). But this second main division of the book treats of some very important matters. Kant's mature philosophy is called the "critical" philosophy because everything he wrote subsequent to 1781 is conceived as resting on the *Critique of Pure Reason.* Arguably more is said directly about this in the Doctrine of Method than in the earlier, more frequently studied parts of the *Critique.*

The discipline of pure reason

Human reason, in its theoretical use, has been shown to be confined within narrow limits. Its primary task, a difficult one, is to discipline itself in light of this self-knowledge. Kant divides this discipline into four sections: (i) the "dogmatic" use of reason, (ii) the "polemical" use of reason, (iii) reason's hypotheses, and (iv) reason's proofs.

The first section includes Kant's most complete discussion anywhere of the science of mathematics (KrV A712–738/B740–766). Its aim is to argue against the attempt (found in the philosophies of Descartes, Spinoza, and Leibniz) to imitate the method of mathematics in other branches of philosophy (especially in metaphysics). Mathematics, he argues, has certain distinctive advantages over other sciences, owing to its inherent limitation to what can be exhibited *a priori* in the pure intuitions of space and time. Properly speaking, it is only in mathematics that we can find genuine definitions, axioms, or demonstrations. When philosophers present their theories as if they could avail themselves of these features of mathematics, they only deceive themselves, presenting arbitrary concepts and (necessarily groundless) inventions as if they could have the same kind of necessary and non-empirical grounding appropriate to mathematical theorems. The third and fourth sections further develop this critique on the methods of metaphysics by prescribing limits on what reason should employ as hypotheses or offer as proofs in matters that transcend empirical inquiry.

In the second section, dealing with the "polemical" use of reason, Kant turns to an ardent defense of freedom of public communication, and of a spirit of open-mindedness in the discussion of metaphysical issues, arguing that the very existence of reason itself depends on the free give and take of controversy between rational beings, which requires the liberty to come to one's own conclusions honestly and express them openly to others (KrV A738–769/B766–797). This discussion is distinctive in that it directly connects the concerns of theoretical reason or science with considerations that are moral or political in nature. Chief among Kant's concerns here are to protect the liberty of

thought and its expression against political repression that is motivated by *religious* concerns, and which regards all critical questioning of religious dogmas as morally or spiritually harmful either to the individual soul or the political order. Kant was to return to these issues many times in his later writings, especially in the essays *An Answer to the Question: What is Enlightenment?* (1784, Ak 8:35–42), *What Does It Mean to Orient Oneself in Thinking?* (1786, Ak 8:133–146), and *Conflict of the Faculties* (1798, Ak 7:5–116).

The canon of pure reason

In the second chapter of the Doctrine of Method Kant argues that reason's pretensions must be limited, but they cannot be checked externally, by censorship; instead, they must be checked internally, by reason itself, which therefore requires a 'canon', or set of principles determining how it should form its beliefs. Kant's principal thesis here is that reason requires such a canon not from a theoretical standpoint but only from a practical or moral standpoint, so that in matters that transcend its theoretical capacities, the propositions it holds to be true may be consistent with the moral duties reason prescribes to itself. The Canon of Pure Reason includes not only Kant's first systematic statements of his argument for rational faith in God on moral grounds (which we will discuss further in chapter 9), but also his most systematic discussion of moral philosophy prior to the *Groundwork for the Metaphysics of Morals* (1785), which will be the topic of chapter 7.

The architectonic and history of pure reason

In the concluding two sections of the Doctrine of Method, the 'Architectonic of Pure Reason' and the 'History of Pure Reason', Kant attempts to outline the entire system of philosophical knowledge in light of the findings of the *Critique.* We therefore learn how Kant's other main works are related to the system of philosophy he is attempting to found. The History of Pure Reason, for all its tantalizing brevity, is an attempt by Kant to conclude by orienting the critical philosophy clearly in relation to the positions (dogmatism, empiricism, skepticism, indifferentism) he discussed metaphorically in the Preface to the first edition.

The basic theme of the *Critique of Pure Reason* is the limitedness of reason. No philosopher has laid more stress than Kant did on the importance for human beings of keeping in mind the limited capacity of their reason in all the affairs of life, especially in the conduct of inquiry and formation of beliefs. Yet no philosopher ever asserted more ardently the absolute title of reason to govern human thought and action, or gave us sterner warnings concerning the inherent badness and the disastrous

consequences of permitting human passions, enthusiasms, or inspirations, or the supernatural deliverances of authority or tradition, to usurp the authority of reason. The *Critique* gives the lie to all those who, standing in the Romantic tradition, assert that Enlightenment rationalism errs in overestimating our rational capacities or being insufficiently attentive to their limitations. On the contrary, the truly dangerous error is to imagine that human beings have access to some faculty or source of wisdom higher than reason, exempt from rational criticism, and to be followed in preference to it. The importance of subjecting reason itself to critique lies precisely in the fact that beyond reason there can be no legitimate appeal.

further reading

Henry Allison, *Kant's Theory of Freedom*. New York: Cambridge University Press, 1990.

Karl Ameriks, *Kant's Theory of Mind*. Oxford: Clarendon Press, 1982.

Jonathan Bennett, *Kant's Dialectic*. Cambridge, England: Cambridge University Press, 1974.

Michelle Grier, *Kant's Theory of Transcendental Illusion*. Cambridge, England: Cambridge University Press, 2001.

Allen W. Wood, *Kant's Rational Theology*. Ithaca: Cornell University Press, 1978.

— (ed.), *Self and Nature in Kant's Philosophy*. Ithaca: Cornell University Press, 1984.

notes

1 See Paul Guyer, *Kant and the Claims of Knowledge* (New York: Cambridge University Press, 1987), pp. 385–404.

2 For a more detailed discussion of this topic, see "Kant's Compatibilism," in Allen Wood (ed.), *Self and Nature in Kant's Philosophy* (Ithaca: Cornell University Press, 1984), pp. 57–72.

3 This way of reading Kant is well known from the writings of Christine Korsgaard, but probably its most resolute development is found in Hilary Bok, *Freedom and Responsibility* (Princeton: Princeton University Press, 1998).

chapter 6

philosophy of history

Kant's writings on human history appear at first glance to constitute only a small part of his literary output and to have only marginal significance for his philosophy. Unlike some other great modern philosophers, such as Leibniz, Hume, and Hegel, Kant was not himself a historian, not even a very well read historian of philosophy. The essays devoted chiefly to the philosophy of history consist in a few brief occasional pieces, such as *Idea for a Universal History with a Cosmopolitan Aim* (1784) and *Conjectural Beginning of Human History* (1786), plus some parts of other essays, such as the one about the common saying on theory and practice (1794) or the *Conflict of the Faculties* (1798). But if we look more closely at some of his most important works, we begin to see that views about history, even quite distinctively Kantian views, play a major role in their arguments and even in their very conception.

Probably Kant's most conspicuous appeal to his philosophy of history occurs in the "First Supplement" in *Toward Perpetual Peace*, in the form of the 'guarantee' he offers for the terms of peace between nation states that he has proposed (EF 8:360–368). But we have seen that the Prefaces to both the first and second editions of the *Critique of Pure Reason* reveal that Kant frames his very conception of that work in terms of the *history* of metaphysics as a science. Reflections on the philosophy of history also play a role in the argument of the closing pages of the *Critique of the Power of Judgment*, where Kant is attempting to bridge the gulf between theoretical understanding and practical reason by relating the ultimate end of nature (regarded theoretically as a teleological system) to the final end set by morality (KU 5:429–434). In *Religion Within the Boundaries of Mere Reason*, the history of religion is a prominent part of Kant's expression of hope for the moral progress of humanity (R 6:124–137). And Kant's *Anthropology from a Pragmatic Standpoint* concludes with reflections on the history of the human species (VA 7:321–333). Indeed, Kant's basic characterization of the human species in terms of its collective possibilities for rational self-direction indicates that his conception of human nature itself is a historical conception.

Human history is first of all a collection of facts about what human beings have done and undergone, in which rational human inquiry needs to find some kind of intelligibility. But history seems to be made up of merely contingent facts about the arbitrary actions and accidental good or bad fortune of individuals. (In Voltaire's words, "history is little else than a picture of human crimes and misfortunes.") There seems no guarantee in advance that as a whole, or even in any significant parts, it should be rationally comprehensible at all. But as rational inquirers, we necessarily, and rightly, seek intelligibility in it. Our need to find history intelligible, moreover, is inevitably bound up with a practical interest in it. We hope to find history intelligible in order to make our own actions intelligible to ourselves insofar as they constitute a part of history, perhaps also in order to direct our actions in accordance with historical trends or movements because of the intelligibility they have, especially because of the way our actions may fulfill possibilities or purposes we discover in history. Kant's philosophy of history is fundamentally guided first by the concern to discover something rationally comprehensible in the seemingly accidental occurrences that make up history, and second by the need to relate that understanding to our practical concerns and hopes. In understanding Kant's philosophy of history, it is especially important to recognize the distinctness of these two guiding threads, and the necessary independence of the first from the second.

In Kant's writings about history it is especially conspicuous that his project of understanding human history is bound up with certain rational aims and hopes – for the growth of enlightenment, moral progress of the human species, perpetual peace between nations. These hopes are sometimes related by Kant himself to religious hopes, as when he describes the hope for perpetual peace by saying that "philosophy can also have its 'chiliasm'" – its millenarian expectations (I 8:27). It is therefore not uncommon for expositions of Kant's philosophy of history to interpret this entire philosophy as motivated by practical considerations and consisting in large part of rational hopes analogous to his "practical postulates" of God, freedom, and immortality. Kant's theory of history is then seen as basically an expression of moral-religious hope rather than a program for empirical-factual inquiry. On this interpretation, it consists not of a theory about facts grounded on evidence but in a kind of religious faith grounded *a priori* on moral duties and ends.

There is no doubt that Kant sometimes looked at history in light of our moral vocation and the moral-religious hopes grounded on it. This approach is particularly prominent in his reply to Moses Mendelssohn's rejection of the idea of moral progress in history, found in the third part of Kant's essay on theory and practice (TP 8:307–313). Yet such a reading of Kant's philosophy of history as a whole, and especially of the project set forth in Kant's chief and basic work on the subject – the *Idea for a*

Universal History – is fundamentally mistaken, even a gross distortion of Kant's views about the way human history should be studied and comprehended. In fact, Kant is concerned to *reconcile* and *integrate* a purely *theoretical* concern with making intelligible the welter of contingent facts of which human history consists, with our inevitable and proper concern about the course of history as historical beings and moral agents. This project of reconciliation is quite subtle, and also presents us with a model for the many-sided approach to history found in the great nineteenth-century theorists of history in the German Idealist tradition, Fichte, Hegel, and Marx. But if we do not recognize the independence of the purely theoretical project from the practical concern with it, the entire Kantian project of reconciliation, and even the need for it, will be invisible to us.

A close look at the text of *Idea for a Universal History* reveals that Kant's starting point for the philosophy of history in general is purely theoretical. He does not introduce considerations of a moral-religious nature until the Ninth (and last) proposition of that essay. The right way to describe his approach is to say that he proceeds from considerations of theoretical reason, projecting the "idea" (or *a priori* rational concept) of a purely theoretical program for making comprehensible sense of the accidental facts of human history. He then attempts to bring history as a theoretical object of study, so conceived, into a kind of *convergence* with our practical concerns, so as to unite our theoretical understanding of history with our moral-religious hopes as historical beings. Thus although this essay was written six years before the *Critique of the Power of Judgment*, it already exhibits Kant's attempt in the Methodology of Teleological Judgment to bridge the gulf between theoretical and practical reason. But it can do this regarding history only if it begins by studying history from a purely theoretical standpoint, since otherwise there would be nothing with which to bring our practical hopes into convergence.

The attempt to read Kant's entire philosophy of history as exclusively, or even fundamentally, an exercise in practical faith not only conflicts with the text of *Idea for a Universal History*, but considered in the context of Kant's philosophy as a whole it makes no sense. For Kant the ideas of God, freedom, and immortality are proper objects of rational faith only because these objects, if they have any, would be transcendent to any possible experience, and hence it would be theoretically undecidable in principle whether such objects exist. It is only in the case of such theoretically problematic objects that moral faith is permitted to decide the question (KrV A828–829/B856–857). But human history is a domain within the empirical world, and the undecidability of any beliefs or hopes we might have about it is due not to the fact that no experience is relevant to them, but to the fact that the evidence is too complex and

confused to permit us any firm conclusions. In Kant's view, it is impermissible, even intellectually dishonest, to appeal to practical faith to decide dubious matters of empirical fact.

Further, the historical ends regarding which we might have practically grounded hope – for instance, perpetual peace between nations – are not (like the pure ideal of the highest good) set *a priori* by reason. They are formed through the application of *a priori* practical principles to the empirical conditions of human life. The setting of these ends thus depends in part on propositions about history that must be arrived at through the working out of Kant's philosophy of history. It would be incoherent, or at least question-begging, to attempt to base the philosophy of history itself solely on beliefs held purely on practical grounds where the ends grounding the practical justification of these beliefs must depend in part on that very philosophy of history itself. Practically grounded hopes and beliefs regarding history make sense only relative to a prior and independent theoretical understanding of history and the practical possibilities it affords the human species. It is only on the basis of such a theoretical understanding that we might formulate ends for whose attainment we might have moral grounds to hope. Kant's philosophy of history, as outlined in the *Idea for a Universal History*, aims firstly and chiefly at that purely theoretical understanding.

Natural teleology and human history

The "idea" referred to in the title of Kant's essay is the conception of a *theoretical* project whose aim is to ground the empirical inquiry into human history. It is an "idea" because it is a concept devised starting with *a priori* regulative principles of reason. More specifically, it is devised in accordance with Kant's theory of natural teleology (of which he did not give a full account until the *Critique of the Power of Judgment*, six years later), and in particular in accordance with Kant's conception of the natural teleology of human beings regarded as an animal species.

Kant begins *Idea for a Universal History* by reflecting on the fact that human history is a realm of empirical contingencies, of which however rational inquiry has the task of making sense according to regularities of some kind. As the chief source of this contingency he cites human freedom, which releases people from the regularities of animal instinct, but (so far, at any rate) subjects their actions to no conscious collective rational plan (I 8:17–18). On the basis of Kant's solution to the metaphysical problem of free will in the first and second *Critiques*, it is sometimes thought that he regards human actions in the phenomenal world as capable of being brought under necessary causal laws and investigated like the motions of the heavens or other physical phenomena (see KrV

A550/B578, KpV 5:99), while freedom belongs entirely to the noumenal self. But this is a misreading of (or rather a fallacious inference from) Kant's solution to the problem of freedom. Kant does think that the metaphysical problem of freedom can be solved only by postulating a transcendentally free cause in the noumenal world, and he does think that human actions in the phenomenal world are not exempted from natural necessity. But it does not follow that the species of natural necessity governing human volition is knowable by us, and in fact Kant thinks it is not knowable. Regarding our actions, the future is therefore "not discoverable from known laws of nature (as with eclipses of the sun and moon, which can be foretold by natural means" (SF 7:79). Kant's entire "pragmatic" approach to anthropology (the study of human nature) was predicated on his rejection in the early 1770s of Ernst Platner's "physiological" approach to the subject (see Ak 10:146). For Kant it is an empirical sign (though not a proof) of our freedom that our volitions are not governed by instinct or physiological laws or other discoverable natural regularities.

History and biology

Kant regards living organisms generally as beings whose arrangement and behavior exhibit conceptualizable regularities that cannot be explained by being brought under the kinds of (mechanical) causal laws that make physical phenomena intelligible to us. Instead, they can be brought under the regulative concept of an "organized being" – a being whose internal arrangements and behavior produces its own organic form, and which can therefore be described as "both cause and effect of itself" (KU 5:370). No being in nature corresponds perfectly to this concept, but there are beings in nature (living organisms) that approximate to it, and the investigation of their life processes is governed by a set of regulative principles or maxims, amounting to the assumption (which is not to be taken dogmatically, but used only heuristically) that in an organized being "everything is an end and reciprocally a means as well" (KU 5:376) – in other words, that the life processes of the organized being *maximize* the teleological intelligibility we are looking for. The rationale for this assumption is that we have everything to gain by assuming maximal teleological interconnection in organized beings, since this will guide us toward discovering whatever teleology is present there, and the absence of teleology represents only an empirical limit to the intelligibility of the organism for us, so that there is no cognitive gain to us in ever being satisfied that teleology is absent.

"Since human beings in their endeavors do not behave on the whole merely instinctively, like animals, and yet also not like rational citizens of the world, in accordance with a common plan, no history of them in

conformity to a plan (as perhaps of bees or of beavers) appears to be possible" (I 8:17). Yet there are observable regularities among the free actions of human beings as regards their effects.

> Marriages, the births that come from them and deaths, since the free will of human beings has so great an influence on them, seem to be subject to no rule in accordance with which one could determine their number through calculation; and yet the annual tables of them in large countries prove that they happen in accordance with constant laws of nature, just as much as the inconstant storms, whose single occurrence one cannot previously determine, but which on the whole do not fail to sustain the growth of plants, the course of streams and other natural arrangements in a uniform uninterrupted course. (I 8:17)

Kant's philosophy of history depends on attributing a natural teleology, or unconscious, unintended goal-directedness, to historical events. Because the facts to be made sense of involve the behavior over long periods of time of many human individuals, the natural teleology in history must involve ends that direct the *collective* actions of many human beings, in fact, of many generations of human beings. But because human beings do not co-ordinate their actions "like rational citizens of the world, in accordance with a common plan," this purposiveness must be unconscious, unintended; it must be a *natural* purposiveness, like that found in the organic arrangement of plants and animals. Kant's idea for a universal history is a regulative idea for the investigation of history, guided by the heuristic assumption that human history is guided by a natural teleology.

Since humanity is a species of living organisms, Kant looks for a natural teleology in history in connection with the natural teleology we discover in human beings as living organisms. One heuristic assumption we employ in the investigation of organisms has to do with the development of individual specimens to maturity. It involves the conception of a natural "predisposition" – a global tendency of the organism to develop the set of capacities best suited to carrying on its mode of life. The regulative maxim governing the investigation of predispositions is: "All natural predispositions of a creature are determined sometime to develop themselves completely and purposively" (I 8:18). That is, on purely methodological grounds, we count something as a natural predisposition only if, in the normal and unhindered development of the organism, it develops completely and suitably to the life processes of the species. And in investigating the growth processes of an organism, we conceptualize the global tendencies that show themselves in these processes (such as the development of the capacity to hunt, or find a mate) around the full development of such predispositions. A predatory animal, for example, develops predispositions enabling it to stalk and kill its prey,

while the herbivorous prey animal develops predispositions enabling it to hide from, flee, or repel predators, as well as predispositions enabling it to find and eat the kinds of plants on which it lives.

Kant's First Proposition in the *Idea for a Universal History* invokes this teleological maxim, and then the Second Proposition applies it, in an extended and creative way, to the human species, in light of its distinctive capacities as a species of free and rational beings. Reason is a capacity that frees those beings that have it from the limitation to only one way of life, and enables them to invent, so to speak, their own nature and their role in the natural world (MA 8:111–112). It gives human beings what Rousseau called "perfectibility."[1] The predispositions of rational beings, therefore, are not fixed by instinct, as they are for other animals, but devised by human beings themselves. From this follows Kant's Third Proposition: "Nature has willed that the human being should produce everything that goes beyond the mechanical arrangement of his animal existence entirely out of himself, and participate in no other happiness or perfection than that which he has procured for himself free from instinct through his own reason" (I 8:19).

The economic basis of history

Further, it implies that human predispositions are handed down from one generation of human beings to the next, and then they are modified or augmented by the reason of those who receive them. Consequently, what we count as the predispositions of the human species are continually developing and growing, and the heuristic maxim that nature has ordered things in such a way that all of them eventually develop fully amounts to the claim that human history exhibits a tendency, unintended by human beings themselves, toward the accumulation and boundless development of human faculties and diverse ways of life, with those ways of life predominating that enable these faculties to be exercised to the full and to develop further. In his essay *Conjectural Beginning of Human History*, Kant distinguishes different phases or stages of human history, based on the historically developed way of life that is dominant in them. In the first phase, people lived as hunter-gatherers; in the next, they tamed animals and lived a pastoral life as nomadic herders (MA 8:118–119).

Then, according to Kant's account, came the true revolution in human history, when people developed the capacity to plant and grow crops. Agriculture necessitated a settled mode of life, both in order to reap the harvests they sowed and in order to live off the stored-up products. It required that the producers limit themselves to certain parts of the earth's surface, but also that they defend those parts against others, and in particular against the incursions of those who still practiced the more

primitive ways of life, such as the herdsmen who wanted to drive their flocks across the cultivated land. Farming was the most productive mode of life devised so far, creating a surplus, teaching people to plan their lives and defer the satisfaction of their needs, and freeing them to diversify their activities. This led to the creation of towns and the development of diverse practical arts, and a division of labor. Part of the productive surplus could be, and had to be, devoted to the creation of the coercive force needed to protect the rights of property, in both land and stored-up goods, that made the agricultural and urban ways of life possible (MA 8:119–120). The protection of property, for Kant, as for Locke, Rousseau, and many other modern political theorists, represents the fundamental rationale and function of civil society and the foundation of all those legitimate coercive institutions concerned with the protection of rights and justice.

Not to be missed here is the way in which Kant's philosophy of history on these points anticipates the Marxian materialist conception of history. Marx too sees history as divided into stages that are characterized by modes of production fundamentally distinguished by the degree of development of the productive forces of society. And he too sees political institutions as based on the property relations corresponding to the prevailing mode of production. Kant's theory of course lacks the Marxian conception of class conflict as the determinant of social dynamics, but Kant does view social change as involving conflict between higher and lower productive modes – such as the conflict between the pastoral and agricultural ways of life.

Unsociable sociability

In another way too, however, Kant's philosophy of history also grounds social progress just as deeply on social conflict. For in his Fourth Proposition in the *Idea for a Universal History*, Kant identifies the mechanism through which he thinks human predispositions unfold in history. This mechanism is social antagonism, a propensity in human nature to compete with other human beings, to have one's own way against the will of others, and to achieve superior rank or status in the opinion of others. Alluding to a remark by Montaigne (one of Kant's favorite authors), he calls this propensity of human nature "unsociable sociability"[2] – meaning that it is simultaneously a propensity to be dependent on others (for one's sense of superiority to them) and also a propensity to cross others, to isolate oneself from them, and to behave unsociably within this fundamental relation of interdependency. Through unsociable sociability, we seek honor, power, and wealth, that is, superiority to others exercised over them (respectively) through their opinion, their fear, or their interest. These are the three objects of the

social passions (VA 7:271–275), that is, inclinations that are difficult for us to control through reason. In the *Critique of Practical Reason*, unsociable sociability appears as "self-conceit" (KpV 5:72); in *Religion Within the Boundaries of Mere Reason*, it appears once again as the radical propensity to evil in human nature (R 6:29–44).

Unsociable sociability develops along with the same faculty of reason that enables us to know that it is evil; both are products of society. Acting from our propensity to unsociable sociability is something we do freely, and for which we are to blame. But there is natural purposiveness in unsociable sociability – in other words, nature employs this propensity to further the development of the predispositions of the human species. When people seek to gain superiority over others, they make themselves both unhappy and evil. But in the process they develop capacities that are passed along to later generations and enrich both human nature and human history.

The political state

As a mechanism for developing human predispositions, however, unsociable sociability reaches a limit at the point where human conflict disrupts the stable life of civilization that is needed for the preservation and further development of human faculties. If life and property become insecure, then people have no opportunity to perfect themselves and no incentive to accumulate products of labor, which may be taken from them before they can be enjoyed. At a certain point, therefore, nature's end of endlessly developing the predispositions of the human species requires a stable and ordered society, a condition of peace with justice. When civilization reaches this point, natural purposiveness requires another device alongside unsociable sociability to balance its counterpurposive effects. This device, which Kant introduces in the Fifth and Sixth Propositions of *Idea for a Universal History*, is the establishment of "a universal civil society administering universal right" (I 8:22). This civil society, characterized by a coercive power protecting rights and property, is the political state. It is a voluntary creation of human beings themselves, and is subject to ideal rational principles (of right or justice) that people are capable of recognizing and obeying; but in promoting the full development of our species predispositions, the establishment of a political state also accords with natural teleology.

The creation of a perfect civil constitution for the state presents itself to the human species as a "problem" to be solved by people themselves, because in addition to being an end of nature (needed to facilitate nature's more basic end of developing human faculties), justice among human beings also presents itself to them as a demand of reason, something they unconditionally ought to achieve. In the Seventh Proposition,

Kant argues that this problem cannot be solved as long as states remain in a permanent state of war in relation to one another. For not only are wars themselves destructive of the conditions needed to develop human faculties, but the continuous need to be prepared for war distorts the state by putting power in the hands of those who would govern in the spirit of military despotism and by diverting human talents and resources to aims irrelevant or hostile to human progress.

Kant is not ignorant of the arguments, advanced in his own century by Turgot and renewed in the twentieth century by partisans of the Cold War, that military technologies too can serve human progress.[3] Nor is he unsusceptible to the idea, most often associated with Hegel, that the sublimity of war helps to unite the state and raise individuals above the ignoble disposition to complacent self-seeking that characterizes private economic life in peacetime (indeed, Kant even directly expresses this "Hegelian" idea himself, KU 5:263).[4] But Kant thinks that the stage of history in which armed conflict between states, and the preparations for conflict, are conducive to human progress, is a cruder one than that which the human species has now attained, at least in civilized parts of the world.

The theoretical comprehension of history and moral striving

The result is, as Kant puts it in the Eighth Proposition, that both progress toward a perfect state constitution, and the creation of a peaceful international order among states, may be regarded as ends of nature in history. "One can regard the history of the human species in the large as the completion of a hidden plan of nature to bring about an inwardly and, to this end, also an externally perfect state constitution, as the single condition in which nature can fully develop all its predispositions in humanity" (I 8:27).

It is crucial to distinguish between two quite different (and largely independent) theses here. The first thesis, which is the primary focus of *Idea for a Universal History*, is a wholly theoretical one: Under the guidance of heuristic or regulative principles of reason, we should attempt to make sense of human history as a process involving an unconscious and unintended teleology of nature, whose ultimate end regarding the human species is the open-ended development of its predispositions, and whose ends, subordinate to this one, also include the creation of a perfectly just civil constitution and a peaceful international order among states. A second thesis, evident at many points in Kant's writings, extremely important to his philosophy as a whole but of only ancillary significance in the *Idea for a Universal History*, is a practical or moral one: As human beings, we have a duty to work together toward devising and realizing the end of a perfect civil constitution administering

justice among human beings, and to this end we are also required to seek an order guaranteeing perpetual peace among states.

The first thesis has no moral presuppositions. It results in part from *a priori* regulative principles of reason when they are applied to the facts of human history, and in part from these facts themselves, such as the fact that nature is seen to employ unsociable sociability as the device for unfolding the predispositions of the human species, and the fact that beyond a certain point this device can continue to operate toward nature's end only if it is counterbalanced by a humanly created order of peace with justice within the political state and between states. The second thesis is a purely practical (or moral) one, deriving from the fact that human beings, as rational beings, are ends in themselves, consequently beings whose external freedom ought to be protected, and whose perfection and happiness ought to be set as ends by all rational beings.

The first (theoretical) thesis is in no way dependent on the second (practical or moral) thesis. For Kant it is practical reason, not natural purposiveness, that grounds our moral duties. If something is an end of morality for practical purposes, it does not follow that it should be regarded by theoretical reason as an end of nature. Nor does the fact that something should be treated for heuristic purposes as an end of nature necessarily imply that there is any moral reason to promote it. Kant does think that some of our duties (for instance, our duties to ourselves regarding self-preservation, and the use of food, drink, and sex) derive from respecting the natural purposiveness of our organization as living beings. But unsociable sociability also introduces a natural purposiveness into our lives, which inclines us to seek superiority – through honor, wealth, and tyrannical dominion – over other human beings, who are our equals in the eyes of reason, and to treat them as mere means to our own selfish ends. But such conduct is paradigmatic of what violates the moral law, and the fact that such conduct serves natural purposes is no justification or excuse for it. Kant is quite explicit that natural teleology by itself does not entail any moral duty to co-operate with it: "When I say of nature, it *wills* this or that to happen, this does not mean, it lays upon us a duty to do it (for only practical reason, without coercion, can do that) but rather that nature itself *does* it, whether we will it or not (*fata volentem ducunt, nolentem trahunt*)" (EF 8:365).[5] In other words, when we have a duty to do something that accords with nature's ends, natural teleology co-operates with us, but when moral ends oppose natural ends, nature will offer resistance to the good will.

Yet there is one connection between Kant's theoretical theses about history and his practical ones. This connection makes the former thesis part of our grounds for the latter. The fact that, according to a teleologically conceived theoretical philosophy of history, a natural purposiveness leads toward an ideal civil constitution and toward perpetual peace

between nations – and even more, the factual reasons why it does so – constitutes part of the reason why we have a moral duty to place an ideal civil constitution and perpetual peace among the ends of our action. Moral reason, which recognizes human beings as ends in themselves, provides us with a ground for respecting their rights, and valuing arrangements that protect those rights. But it gives us moral grounds for pursuing an ideal civil constitution only under certain contingent, empirical conditions, specifically, under conditions where human beings themselves have established institutions for the protection of human rights through collective coercive action in the form of a civil constitution, and where there exist imperfect forms of civil constitution presenting us with historical possibilities for improving them.

When we look at history as containing a natural purposiveness toward the perfection of civil constitutions, this gives us a moral reason for cooperating with that purposiveness. We have seen that for Kant himself, the very existence of civil constitutions themselves is historically contingent on the emergence of an agricultural way of life, the growth of urban centers, and the productive surplus made possible by these socio-economic forms; he further understands the imperfect state of civil constitutions as resulting from the fact that they arose as military despotisms reflecting the social conditions determined by the unsociable sociability of human beings. The entire context in which it is rational to set the improvement of civil constitutions as a moral end is conditioned by empirical contingencies highlighted by Kant's philosophy of history.

It is also at most a contingent fact that at the present stage of history the further improvement of civil constitutions must depend on achieving peace between nations. (As we have seen, not everyone, either in Kant's day or our own, even thinks this is a fact at all.) It is therefore only contingent historical facts, along with *a priori* moral principles, that give us any reason to seek perpetual peace as part of the process of striving toward a perfect civil constitution. Only Kant's philosophy of history, regarded as a heuristically motivated project for obtaining and systematizing theoretical knowledge about history, could deliver the kinds of information necessary to warrant our setting perpetual peace between nations and a perfect civil constitution as ends of morality.

History and moral faith

This entails that it could make no sense to view Kant's philosophy of history itself as motivated by moral faith in, or hope for, the achievement of these historical ends. For, as we said earlier, the rationale for setting these ends themselves is not wholly *a priori*, but it depends on theoretical conclusions of fact that, in the context of Kant's philosophy

of history, could be motivated only by the results of his project of theoretical inquiry. To interpret Kant as reasoning in such a manner solely from moral aspirations to historical conclusions is not only to caricature Kantian moral faith as nothing but groundless wishful thinking, but it is also to display the moral hopes themselves as grounded on ends that are without adequate rational motivation in the first place. Such an interpretation, therefore, not only fails to correspond to the letter of Kant's texts on the philosophy of history, but it is also a conspicuously hostile interpretation, which, if correct, could only invite us to dismiss Kant's entire philosophy of history as rationally unmotivated and not to be taken seriously. Kant's philosophy of history makes sense at all only if, in the words of one recent writer on this topic, we see it as satisfying *both* a theoretical and a practical need of reason.[6] Further, we have to see the mode of satisfaction of the theoretical need as coming prior in the rational order even to the emergence of the practical need.

The temptation to think that the theoretical principles of Kant's philosophy of history are really motivated by practical considerations may point to a real problem, however. It seems to be only a lucky contingency that as we come to comprehend human history theoretically, according to regulative principles of teleological judgment, there emerge a set of practical (moral) goals or ends – the perfection of the constitution of civil society, perpetual peace between nations, the moral improvement of the human species. We may find this coincidence suspicious. Why should the best theoretical account of our history point the way to a moral task (or even a series of tasks), perhaps also giving us reason to hope that they can be accomplished? One possible way to disarm this suspicion is to attribute to Kant the view that the entire enterprise has been motivated from the start by moral hopes, so the suspicion itself rests on a misunderstanding of Kant's aims. But I have been arguing that this way out does not represent a tenable interpretation of what Kant says and does in the philosophy of history. If there is a relation of dependence, it seems to go in the opposite direction – the theoretical project of history helps us to understand what specific historical ends moral reason should set in accordance with its *a priori* principles.

It may still seem a suspicious coincidence, however, that a rational comprehension of history should even make this possible. If there is a general answer to such suspicions, I think it would have to consist in Kant's attempt (which increasingly preoccupied him in the second and third *Critiques*) to "reconcile" theoretical with practical reason, or demonstrate their "unity." Kant himself, in other words, also felt perplexed about the relation between the theoretical and practical sides of his philosophy, not only in the area of history, but also in our knowledge of, and action in, nature as a whole, in the gulf between unanswerable metaphysical questions and our religious need for answers to them,

and in the relation between our cognition of nature and our aesthetic responses to it. He wanted to make all these connections more intelligible, to unify under principles what looks disturbingly like a fortunate (or a suspicious) coincidence. Yet it is far from clear what Kant intended to achieve under these headings, and perhaps even less clear whether he ever achieved it. We may just have to learn to live with our disquieting suspicions.

We should also not take it for granted that Kant's philosophy of history is supposed to give us reasons for *expecting* or *predicting* the success of our moral strivings – the actual progress of civil constitutions toward perfection, the actual co-operation of states in a lawful federation maintaining perpetual peace between them. At times Kant appears to think that it does provide such reasons, for in *Perpetual Peace* he offers conclusions from the philosophy of history as providing a "guarantee" of the terms of perpetual peace he has outlined (EF 8:360–368). He realizes that before the skeptical (sometimes cynical, often fearful) heads of state to whom he is addressing his treatise are going to take steps to bring about a peaceful federation, they will need reassurances that the course to which Kant is directing them has some prospect of success in human history. Yet it is not immediately clear that his philosophy of history can offer them these assurances while remaining consistent with its own theoretical claims. For Kant, to identify something as an end of nature is to say that we have heuristic or regulative reasons to look at the facts on the assumption that there are natural tendencies at work to actualize it. But these heuristic reasons by themselves provide us with no theoretical guarantee, no real *evidence* at all, in fact, that what has been identified as a natural end will actually come about. The heuristic recommendation says only that we maximize intelligibility by *looking for* such evidence; it emphatically does *not* say that we are guaranteed to find what we are looking for.

Yet it would not make empirical sense, even for heuristic purposes, to identify something as a natural end if we could not observe some mechanisms at work toward achieving it. We say that maintaining a constant body temperature is an end of nature in animals because we notice instinctive behaviors and mechanisms that tend toward increasing their body heat when they are too cold and toward losing body heat when they are too warm. Likewise, it is a necessary condition for viewing the achievement of perpetual peace among nations as an end of nature that we should find some mechanisms at work that tend in that direction. Along these lines, Kant cites the fact that nations can be strong militarily only if they are strong economically, and that the more civilized a nation becomes, the more its economic strength depends on peaceful prosperity. Those nations that do not value peaceful relations with their neighbors, therefore, should increasingly be unable to make war

successfully on them, while those nations that are in the best position to defend themselves should also be the ones most ready to join in a peaceful federation (see EF 8:368, I 8:27–28).

But Kant also appears to acknowledge that the heuristic reasons provided by his theoretical philosophy of history for expecting the success of his project of perpetual peace fall short of providing a genuine theoretical guarantee. It is at this point that he falls back on the moral duty we have to promote the end of perpetual peace, and to hope, on practical or rational-religious grounds, that the end will be achieved. It is this rational hope, more than any theoretical expectation, that he emphasizes as the "guarantee" of perpetual peace (EF 8:360–362). But he apparently concedes that even that hope would be irrational if there were no theoretical grounds at all for expecting it to succeed. And he therefore offers the admittedly less than conclusive combination of heuristic expectations and empirical reasons supporting them, as sufficient to constitute those grounds: "Nature guarantees perpetual peace through the mechanism of human inclinations itself, with an assurance that is admittedly not adequate for predicting its future (theoretically) but that is still enough for practical purposes and makes it a duty to work toward this (not merely chimerical) end" (EF 8:368).

Critical assessment of Kant's philosophy of history

Even when Kant's views are correctly and sympathetically interpreted, how seriously do they deserve to be taken?

Kant's philosophy of history depends on postulating, at least for heuristic purposes, a natural teleology in human history, whose goals are both collective and unconscious. This aspect of Kant's theory may make it seem extravagant, speculative, anti-empirical, and even obscurantist. His philosophy of history in this respect may seem starkly at odds with his reputation for skeptical modesty and epistemic humility and caution. Proponents of so-called 'methodological individualism' will say that it makes sense to appeal to historical tendencies or trends only when their existence can be authenticated and explained in terms of the choices and motivations of individuals, by way of providing 'microfoundations' for them.

Kant is no methodological individualist, but the unconscious collective ends he posits in human history for regulative purposes are not meant to be postulated arbitrarily, and are not supposed to be divorced from empirically observable motivations and actions of individuals. The whole aim of the theory itself is to use these ends to identify those patterns of human motivation and action that have historical efficacy, distinguishing these from the accidental factors in human choice whose relation to history is merely accidental and insignificant. Kant's

method, however, is not to begin with microfoundations and generate historical trends or natural purposes, but to assume certain natural purposes heuristically and use these as a guide to the discovery of the kinds of motives and actions, both individual and collective, that are historically potent.

The natural end of endlessly developing humanity's species-predispositions leads, for instance, to the specification of unsociable sociability (the human traits of discontent and competitiveness) as the basic mechanism for this development; the historical need, at a certain stage of history, for nations to remain at peace in order to perfect their civil constitutions and continue the development of new human capacities, leads us to recognize the importance of commerce and economic prosperity in making nations powerful, and the reluctance of commercially oriented citizens to turn their lives and property over to warlike heads of state to pursue their greedy and barbaric fantasies of military conquest. Kant also attempts to render the natural teleology of history non-arbitrary by linking it to the natural teleology found in human beings as a species of organisms, as he thinks this teleology plays a role in biological investigations.

No doubt there is a measure of theoretical adventurousness in Kant's historical teleology that might unnerve a traditional empiricist. But there is a very analogous (but much less often appreciated) departure from empiricist caution involved in the abstract idealizations used by methodological individualists in constructing their 'microfoundations' for the social trends and tendencies they are willing to countenance. In both cases, an honest assessment of what is going on must take note of the fact that empiricist reconstructions of all domains of knowledge systematically underestimate the creative role of theorizing. There are usually macro-level assumptions built into the choices of abstraction and idealization used to construct microfoundations, and these are all the less subject to empirical constraint and criticism to the extent that they remain unacknowledged. The Kantian choice to begin with the macro-level is more forthright in acknowledging the importance of our *a priori* cognitive ambitions than are the methodological individualists who deceptively pretend they are always sticking close to empirical observation.

Is Kant's philosophy of history outdated?

A more serious problem for Kant's philosophy of history is that we can no longer believe, for instance, in Kant's heuristically motivated natural teleology as the right way for investigating the structure and behavior of living organisms. Since Darwin, it has been recognized that the unconscious and unintended purposive arrangements in living things have a

determinate empirical explanation, based on natural selection. Moreover, this explanation reveals that Kant's heuristic assumption that the teleology in organisms is maximal is empirically and explainably false. When we learn how the organs of a living thing evolved, for instance, we sometimes come to understand why they are not optimally suited for the function they perform. And it might turn out, for similar reasons, that not everything we rightly conceptualize as one of the "species-predispositions" of an organism would have to be fully developed in the normal course of the organism's development. The biological basis of Kant's philosophy of history therefore seems to have been undermined by scientific developments between his time and ours.

Yet it is not so clear that the methodological considerations motivating Kant's philosophy of history are less applicable today than they were in the eighteenth century. Biology may have made advances that undermine the application to it of Kant's heuristically motivated natural teleology, but human history is still an area of inquiry to which no similar empirical theory has been applied with success. It may be that our best chance of making it intelligible is still the regulative-teleological one that Kant adopts. Kant's approach also has the other benefit it had for him, that it enables us to connect an empirical, theoretical study of history to our practical concern with history as historical agents, by identifying historical tendencies (which Kant calls unintended "ends of nature") with which our efforts as moral beings might harmonize. Historical theories since Kant's time (most famously, the historical materialism of Marx) have taken up the idea that historical changes can be understood as functions of the progressive development of collective human capacities, and consequent changes in economic forms over time. Many others besides Marx have used this idea in a wide variety of contexts to deal with social and historical change (for instance, in the many so-called 'theories of modernization'). The basic ideas of Kant's philosophy of history, though they may not always be easily recognizable in their more recent guises, are very far from having been discredited.

It is easier to recognize the way the practical or moral-religious side of Kant's philosophy of history is still with us. We are also still practically concerned with the direction of economic growth and the relation of it to the prospects for peace between nations. Twentieth-century projects for international peace – the League of Nations, the United Nations, and the European Union – are all attempts to fulfill hopes that Kant was among the first to articulate. Also still with us is Kant's cosmopolitanism, which is most fundamentally a view about the historicity of human nature. Kant holds that each of us, while being a citizen of an empirically determinate civil order or political state, is also a citizen of a single world community – our attempt to realize on earth the idea of an

ethical *realm of ends* in which all rational beings are accorded a dignity that is beyond price, and all the ends and maxims should harmonize in one systematic combination. This side of Kant's aspiration is still with us not in the sense that we are very much closer to actualizing the idea, but rather in the sense that it is much more difficult today than it was two centuries ago for a thinking person not to share Kant's sense that we fulfill our nature as human beings only insofar as our species makes historical progress toward it. This in fact is precisely the thought with which Kant closes his last work: *Anthropology from a Pragmatic Standpoint*.

In concluding his attempt to identify the "character of the human species" as a whole, he describes it first in terms of unsociable sociability: "The character of the species, as it is indicated by the experience of all ages and all peoples, is this: that taken collectively (the human race as one whole), it is a multitude of persons, existing successively and side by side, who cannot do without associating peacefully and yet cannot avoid constantly offending one another" (VA 7:331). Then he asks whether such a species should be considered a good race or an evil one, and seems at first to side with those misanthropic critics who either censure humanity for its wickedness, or else laugh at it for its folly – and this not only, he says, through good-natured laughter, but also through a derision of contempt. And these attitudes would be correct, Kant concludes, but for one thing: they themselves reveal in us "a moral predisposition, an innate demand of reason to counteract our evil tendencies." His final conception of human nature, therefore, consists in a *historical* vision of the human species that unites the basis for our criticism with the moral predisposition this criticism reveals.

further reading

Brian Jacobs and Patrick Kain (eds.), *Essays on Kant's Anthropology*. Cambridge, England: Cambridge University Press, 2003.

J. D. McFarland, *Kant's Concept of Teleology*. Edinburgh: Edinburgh University Press, 1970.

Richard Velkley, *Freedom and the End of Reason*. Chicago: University of Chicago Press, 1989.

Yirmiyahu Yovel, *Kant and the Philosophy of History*. Princeton: Princeton University Press, 1980.

notes

1 Rousseau, *Discourse on the Origin of Inequality*, tr. Donald Cress (Indianapolis: Hackett, 1992), p. 26.

2 "Il n'est rien si dissociable et sociable que l'homme: l'un par son vice, l'autre par sa nature." Michel Eyquem de Montaigne, "De la solitude," *Essais*, ed.

André Tournon (Paris: Imprimerie nationale Éditions, 1998), I:388. "There is nothing so unsociable and sociable as man; the one by his vice, the other by his nature," "On Solitude," *Complete Essays*, tr. M. A. Screech (London: Penguin Books, 1991), p. 267.

3 See Anne-Robert-Jacques Turgot, *Turgot on Progress, Sociology and Economics: A Philosophical Review of the Successive Advances of the Human Mind, On Universal History [and] Reflections On the Formation and the Distribution of Wealth*. Translated, edited, and with an introd. by Ronald L. Meek (Cambridge, England: Cambridge University Press, 1973). For a recent defense of this idea in the Cold War context, see Diane R. Kunz, *Guns and Butter: America's Cold War Economic Diplomacy* (New York: The Free Press, 1997).

4 Compare Hegel, *Elements of the Philosophy of Right*, ed. Allen W. Wood, tr. H. B. Nisbet (Cambridge, England: Cambridge University Press, 1991), §324.

5 "The fates lead the willing, drive the unwilling" (Seneca, *Moral Epistles*, 18.4).

6 Pauline Kleingeld, *Fortschritt und Vernunft: Zur Geschichtsphilosophie Kants* (Würzburg: Königshausen & Neumann, 1995), p. 215.

chapter 7

ethical theory

Kant's moral philosophy is grounded on several related values. Its primary idea is that of the rational agent as a self-governing being. This is closely related to the equal dignity of all rational beings as ends in themselves, deserving of respect in all rational actions. These two values are combined in the conception of an ideal community, or "realm of ends," in which every rational being is a legislating member, and in which all the ends of rational beings are to be combined in a single harmonious system as an object of striving by all of them. These basic values, and their philosophical grounding, are articulated in Kant's two principal foundational works in ethics: *Groundwork for the Metaphysics of Morals* (1785) and the Analytic of the *Critique of Practical Reason* (1788).

Kant's direct and acknowledged influence on the history of moral philosophy rests almost exclusively on these two foundational writings in ethics. In Kant's ethical thought, however, these fundamental values are placed in the context of what Kant calls an "empirical anthropology," a distinctive theory of human nature and the human condition. If Kant's theoretical critique is about the limits of reason in its attempt to acquire knowledge *a priori*, then his practical philosophy is about the proper limitations of *empirically conditioned* reason – reason acting in the service of non-rational desires (KpV 5:15–16). The basic Kantian contrast between "duty" and "inclination," and between the *a priori* or "formal" principle of morality and "material" principles based on our natural desires, depend not only on the *a priori* foundations of Kant's theory but also on his theory of human nature. The historical basis of this crucial empirical side of Kant's ethical thought was discussed in the previous chapter. He never developed the "practical anthropology" that he said was needed for a complete moral philosophy (G 4:388), but he did include "anthropological" considerations in the reasonings through which he derived the system of juridical and ethical duties presented in his final work on ethics, the *Metaphysics of Morals* (1797–8).

In addition to the foundational works in ethics and the historical or anthropological writings, Kant also produced a number of writings in

which he applies ethical principles. This includes not only the system of duties in the *Metaphysics of Morals* but also works on politics and religion that constituted his chief output during the last decade in which he wrote: *An Answer to the Question: What Is Enlightenment?* (1784), *What Does It Mean To Orient Oneself in Thinking?* (1786), *On the Common Saying: That May Be Correct in Theory but It Is of No Use in Practice* (1793), *Religion Within the Boundaries of Mere Reason* (1794), *The End of All Things* (1794), *Toward Perpetual Peace* (1795), *On A Presumed Right to Lie from Philanthropy* (1797), and *Conflict of the Faculties* (1798). One result of emphasizing Kant's foundational writings in ethics has been to neglect these writings and to give chief emphasis to Kant's most formalistic statements of the moral principle, and to treat the opposition of the motive of duty or reason to that of feeling or inclination as a consequence of Kant's ethical 'formalism'. In this chapter I will be concerned to correct the misperceptions that have resulted from this misemphasis on Kant's foundational writings in ethics and the neglect of the larger context of writings on anthropology and applied ethics in which they need to be understood. For this reason, I will draw on Kant's philosophy of history, discussed in the last chapter, to provide a context in which the foundations of Kantian ethics should be understood.

Categorical imperatives and inflexible moral rules

But at the outset, some even more elementary sources of misunderstanding and resistance to Kant's ethical theory should be dealt with. Kant notoriously held some very extreme (even repellent) positions on certain ethical issues. He held that murderers should always be put to death, that suicide is contrary to a strict duty to yourself, that sexual intercourse is inherently degrading to our humanity, that masturbation is an even more serious moral crime than suicide, that no disobedience to duly constituted political authority is ever justifiable except when the authority orders you to do something that is in itself wrong, and he once argued that lying for the purpose of adding to human welfare, even to save the life of an innocent person from a would-be murderer, is always wrong.[1] It is not uncommon for unsympathetic interpreters to exaggerate Kant's views on these matters, but even charitably interpreted many of his moral opinions on particular subjects seem inflexible even to the point of inhumanity. Some of these views were idiosyncratic even in his own day, though most were certainly shared in his time far more widely than they are now. If we wish to learn anything from, or about, Kantian moral theory (as distinct from merely providing ourselves with a plausible pretext for refusing to learn from it), then we need to ask about Kant's scandalous opinions whether they actually follow from the values and principles contained in his moral theory.

Perhaps a theory whose fundamental value is the autonomy of reason and the dignity of rational beings can be expected to provide reasons not to accept pleasure and expediency as sufficient grounds to lie or to destroy one's own rational nature. But it is hard to see how such values could justify inflexible rules against lying or suicide, not to mention justifying some of Kant's other scandalous opinions. (Human dignity is also seen as providing reasonable grounds for making exceptions to moral rules against lying or suicide in certain cases.) Those who care about the particular moral issues should look at Kant's own reasoning from his principles to his conclusions, but it should not be taken for granted that such reasoning is valid, or that Kant's views on particular moral issues necessarily represent a correct interpretation of the basic principles of his moral theory.

One way of associating the inflexibility of some of Kant's views with something fundamental to his moral theory is to see them as expressive of his idea that moral duties are "categorical imperatives." Categorical imperatives are supposed to be unconditionally valid. Therefore, any principle that is seen as a categorical imperative (for example, 'Do not lie') must be viewed inflexibly as having no exceptions whatever. But this ridiculously fallacious argument rests on a very simple confusion. For Kant, a rational normative principle (or "imperative") guiding our action is "categorical" if its validity is not conditional on having set some end to which the action is to serve as a means. This does not entail, however, that the validity of rules which, *when they are valid*, are categorical imperatives, cannot be conditional on particular circumstances, or that there cannot be grounds for making exceptions to a generally valid moral rule. When lying is wrong, according to Kant, its wrongness is not conditional on whether some desirable end (such as human happiness) is achieved by abstaining from lies. But it does not follow that there cannot be exceptions to the rule 'Do not lie' – that is, cases in which this rule is not in fact binding as a categorical imperative. How often such exceptions occur must be decided by looking at the derivation of the moral rule 'Do not lie' from more basic Kantian principles, such as "Treat every rational being as an end in itself," and considering possible cases in which this more basic value might not require strict adherence to that rule. Kant treats *exceptivae* (exceptions to moral rules) as one of the twelve fundamental categories of practical reason (KpV 5:66), and the twenty-odd "casuistical questions" that Kant raises about specific duties in the Doctrine of Virtue deal mainly with cases in which there may arguably be exceptions to rules that hold generally, though not universally.

It is true that Kant also regularly calls our attention to (and is highly critical of) the human tendency to make exceptions of ourselves in the case of moral rules we expect others to follow, and to use the fact that

moral rules may sometimes have exceptions as a shabby excuse for failing to follow moral rules when we should follow them. But the passages in which he says those things are surely not open to criticism on the grounds of excessive inflexibility or inhumanity. For Kant is surely right that people do often do this, and that their doing it is responsible for much evil and much that is reprehensible in human conduct.

I practical anthropology

In the *Groundwork for the Metaphysics of Morals*, Kant divides ethics into two parts: the *metaphysics of morals*, consisting in moral principles valid *a priori* for every rational being, and *practical anthropology*, an empirical study of the human nature to which the principles are to be applied (G 4:388). It is too seldom appreciated that Kant there treats practical anthropology as a necessary part of ethics, without which, in his view, it would not be possible to specify determinate duties. Perhaps this is because Kant never wrote a work specifically on practical anthropology, despite the fact that his lectures on anthropology, begun in 1772 and continuing to the end of his teaching career, were the most popular and the most frequently offered lecture course he gave. Kant's various remarks about the present state of our sciences of human nature show him to believe both that despite the importance of this study, there are severe limitations on our capacity to treat it scientifically, and also that the present state of the study of human nature is very poor even in relation to its limited possibilities. It is also less often appreciated than it should be that when he finally came to write a *Metaphysics of Morals* at the very end of his career, Kant recast the distinction between 'metaphysics of morals' and 'practical anthropology', integrating the empirical 'principles of application' into 'metaphysics of morals' itself and restricting 'practical anthropology' to the study of the "subjective conditions in human nature that hinder people or help them in fulfilling the laws of a metaphysics of morals" (MS 6:217).

The only approach to the study of human nature that Kant works out with confidence is to be found in his writings on the philosophy of history. As we saw in the last chapter, Kant's thesis is that human history can be made theoretically intelligible to us only by finding in it a natural end, which is the full (hence temporally endless) development of the natural predispositions of the human species (I 8:18). This end does not belong to the conscious intentions of people, but is a natural end, posited by reflective judgment as a regulative idea for maximizing the intelligibility of the data to us (I 8:17; cf. KU §§75–79, 5:397–417). Since in a rational species, these predispositions do not belong to any individual specimen but only to the entire species as it develops through time, the

ends which make human history intelligible must be collective ends of the whole species through time, which individuals serve unintentionally and of which they can become conscious only through the philosophical study of history (I 8:17–20).

This much already gives us enough to make two points controverting common misunderstandings of Kantian ethics. First, it is not merely oversimplified but fundamentally erroneous to represent Kant as having a 'timeless' or 'ahistorical' conception of reason, and to see Hegel (for example) as "correcting" it by introducing a "historical" conception (this representation badly misreads Hegel too, but there is no time to go into that here). Second, the thesis that human history is grounded on an unconscious collective purposiveness, which is quite rightly associated with German Idealism and more specifically with Hegel, was already fully present in the philosophy of Kant (though for him it was not to be regarded as a dogmatic principle of speculative metaphysics, but a regulative principle of judgment, adopted because it is a necessary heuristic device for making the empirical facts of history intelligible to us).

A third point becomes clear when we look at Kant's execution of his theoretical project in the *Idea for a Universal History*. Human nature develops in history chiefly through competitiveness; each individual seeks to "achieve a rank among his fellows, whom he cannot stand but also cannot leave alone" (I 8:21). The natural history of human reason is therefore a process driven by people's natural inclinations, behind which lurks a propensity to "self-conceit," a desire to be superior to other rational beings, hence to use them as mere means to one's ends and to exempt oneself from general rules one wants others to obey. It is this thesis that grounds Kant's famous (or notorious) suspicion of our empirical desires, or inclinations.

"The human being feels within himself a powerful counterweight to all commands of duty, which reason represents to him as so deserving of the highest respect – the counterweight of his needs and inclinations" (G 4:405). Kant's critics (beginning with Schiller, but including Hegel and countless others down to the present day) read such remarks as the one just quoted in a shallow and shortsighted manner when they attribute it to an artificial metaphysical "dualism," or to an unhealthy (stoical or ascetical) hostility to "nature" or "the senses" or "the body." As Kant makes quite clear, the counterweight to reason and duty is nothing so innocent. The opponent that respect for morality must overcome is always "self-conceit" (KpV 5:73), which arises not out of our animal nature but from our humanity or rationality (R 6:27). The enemy of morality within us is not "to be sought in our natural inclinations, which merely lack discipline and openly display themselves unconcealed to everyone's consciousness, but is rather as it were an invisible enemy, one who hides behind reason and hence is all the more dangerous" (R 6:57).

Our ironic predicament, in Kant's view, is that the natural device of social antagonism is required to develop our rational faculties, which (like all human faculties) belong more to the species than the individual, and show themselves chiefly through our capacity for self-criticism through free communication with others (KrV A xi-xii, A738–739/B766–767, O 8:144–146, KU 5:293–298). When reason develops, however, it recognizes a moral law whose fundamental value is the dignity (or absolute, incomparable worth) of rational nature in every rational being, hence the absolute equality of all rational beings (G 4:428–429, 435, MA 8:114, MS 6:314, 435–437, 462–466). Reason must therefore turn against the very propensity in our nature that made it possible. Kant therefore thinks that the most adequate conception of our human nature that we can form is a historical one, centered on the task of converting ourselves from competitive and antagonistic beings into beings capable of uniting with one another on terms of mutual respect: "What is characteristic of the human species in comparison with the idea of possible rational beings on earth is that nature has put in them the seed of *discord*, and willed that from it their own reason should produce *concord*, or at least the constant approximation to it" (VA 7:322). Our destiny is to be engaged in an endless struggle between "nature" and "culture," whose object is the moral perfection of the human character.

"Natural predispositions, since they were set up in a mere state of nature, suffer violation by progressing culture and also violate it, until perfected art once more becomes nature, which is the ultimate goal of the moral vocation of the human race" (MA 8:117–118). Kant is no more opposed than are his critics to understanding the aim of culture as that of bringing our natural desires into harmony with the demands of reason. His philosophy of history, however, gives him reason to think that this reconciliation will be an extremely long and difficult social process. It is not to be accomplished merely through a philosophical conversion – by the adoption of more "healthy" (that is, more complacent and less self-critical) attitudes toward our desires. Nor will it help to "go beyond dualisms" if that is a euphemism for a state of denial concerning the fact that coming to terms with our nature (especially our corrupt *social* nature) will be an endless, painful historical task.

II the fundamental principle of morality

Kant's aim in the *Groundwork* is to "seek out and establish the fundamental principle of morality" (G 4:392). In the First Section of the *Groundwork*, Kant attempts to derive a formulation of the principle from what he calls "common rational moral cognition," or the

moral know-how he thinks every human being has just in being a rational moral agent. Kant's chief aim here is to distinguish the principle he derives from the kinds of principles that would be favored by moral sense theorists and by those who would base morality on the consequences of actions for human happiness. This attempt is not very successful, because Kant underestimates the extent to which the competing theoretical standpoints are capable of alternative interpretations of the issues and examples he discusses, yielding reactions to them that call into question the responses he regards as self-evident. Thus the opening pages of the *Groundwork*, especially its famous attempt to persuade us that actions have moral worth only when they are done from duty, has seldom won converts to Kant's theory and more often distracted attention from what is really important in Kant's ethical theory. Kant is more successful when he makes a second, more philosophically motivated attempt to expound the moral principle in the Second Section.

Kant thinks that if correct moral judgments are to constitute a well-grounded and consistent whole, they must ultimately be derivable from a single fundamental principle. But in the Second Section of the *Groundwork*, Kant considers this one principle from three different standpoints, and formulates it in three distinct ways. In two of the three cases, he also presents a variant formulation that is supposed to bring that formulation "closer to intuition" and make it easier to apply. The system of formulas can be summarized as follows:

First formula:
FUL *The Formula of Universal Law*: "Act only in accordance with that maxim through which you can at the same time will that it become a universal law" (G 4:421; cf. 4:402);

with its variant,

FLN *The Formula of the Law of Nature*: "Act as if the maxim of your action were to become by your will a universal law of nature" (G 4:421; cf. 4:436).

Second formula:
FH *The Formula of Humanity as End in Itself*: "So act that you use humanity, whether in your own person or that of another, always at the same time as an end, never merely as a means" (G 4:429; cf. 4:436).

Third formula:
FA *Formula of Autonomy*: ". . . the idea of the will of every rational being as a will giving universal law" (G 4:431; cf. 4:432) or "Choose only in such a way that the maxims of your choice are also included as universal law in the same volition" (G 4:439; cf. 4:432, 434, 438).

with its variant,

> FRE *The Formula of the Realm of Ends:* "Act in accordance with the maxims of a universally legislative member of a merely possible realm of ends" (G 4:439; cf. 4:432, 437, 438).

FUL (and FLN) consider the principle of morality merely from the standpoint of its *form*, FH considers it from the standpoint of the *value* which rationally motivates our obedience to it, and FA (and FRE) consider it from the standpoint of the ground of its *authority*.

The Formula of Universal Law

The earliest characterization of Kantian ethics adopted by his German Idealist followers and critics was that Kantian ethics is "formalistic." The use of this epithet is due largely to the mistaken emphasis Kant's readers place on the first formulation of the moral principle at the expense of the other two formulations, whose aim is precisely to complement and hence remedy any such "formalism." From this first standpoint, however, the principle is what Kant calls a "categorical imperative." Kant's terminology here is derived from the logic of his day, but it can mislead us if we are not careful. An *imperative* is any principle through which a rational agent constrains itself to act on the basis of objective grounds or reasons. An imperative is *hypothetical* if the rational constraint is conditional on the agent's adoption of an optional end, and *categorical* if the constraint is not conditional in this way. As long as some hold that all rationality is "only instrumental," it is controversial whether there are (or could be) any categorical imperatives. Kant's procedure in the *Groundwork* is to assume provisionally that there are, and to inquire, in the Second Section, what their principle would have to be. Then in the Third Section Kant attempts to argue that as rational beings we must in effect presuppose that there are such imperatives, which therefore establishes the validity of the formulas derived provisionally in the Second Section.

To say that an imperative is 'categorical' therefore means, once again, only that its bindingness is not conditional on our pursuit of some end we have set independently of it. If there is a categorical imperative to keep promises, this means only that the rational bindingness on us of keeping promises is not conditional on some further end to be achieved through the keeping of promises (such as the self-interested benefits we derive from being able to make contracts with others). But it does not imply that the obligation to keep promises might not be conditional in other ways – for instance, that this obligation might cease to exist if keeping the promise would somehow violate the dignity of humanity or if we knew that the person promised would release us from the promise

if they knew of the unforeseen situation in which we find ourselves when it comes time to keep it. When we have good and sufficient grounds to make exceptions to a moral rule, this means only that the rule (under those circumstances) no longer binds us categorically (or, indeed, in any other way). Thus whether there are any moral rules at all that hold without exceptions is not decided by accepting Kant's claim that all moral obligations involve categorical imperatives.

Because FUL is supposed to be derived from the very idea of a categorical imperative, it is easy to fall into using the term "the Categorical Imperative" simply to refer to it. But this often leads to the unjustifiable privileging of FUL as the principle definitive of Kant's theory, and the consequent neglect of FH and FA. Kant regards his argument in the Second Section of the *Groundwork* as an exposition of the principle of morality, which passes through three stages and reaches completion only at the end of a course of development. This ought to lead us to think of FUL as the starting point of the process. It is the most abstract, most provisional, and (in that sense) the least adequate of the three formulas. And this thought turns out to be right; for it is FH, not FUL, which is Kant's formula of choice for applying the moral principle in the *Metaphysics of Morals*, and it is FA, not FUL, which is used in his attempt to establish the principle in the Third Section of the *Groundwork* (and also in his somewhat different attempt to achieve the same goal in the *Critique of Practical Reason*). The same thought gets confirmed in another way by Kant's critics when, erroneously privileging FUL and virtually excluding FH and FA from their consideration, they then accuse Kant's theory of being satisfied with an "empty formalism." This charge, however, is an indictment less of Kant's theory than of their own shortsighted reading of the *Groundwork*.

FUL is derived from the mere concept of a categorical imperative in the sense that it tells us simply to obey all "universal laws," that is, practical principles that apply necessarily to all rational beings. In order to make this a bit more informative, Kant includes in FUL a test on *maxims* (subjective practical principles, formulating an agent's policies or intentions), which is supposed to determine which maxims conform to universal laws. FUL says that a maxim violates a universal law if it cannot be willed as a universal law. FLN tries to bring this test closer to intuition by inviting us to imagine a system of nature of whose laws the maxim is one, and asking us whether we can, without contradiction or conflicting volitions, will to be a part of such a system of nature. After deriving FUL and FLN, Kant attempts (I think prematurely and over-anxiously) to illustrate his moral principle by applying these tests to four maxims. The maxims are chosen to be typical of the way an agent might be tempted to violate a duty, and the four duties are selected according to a taxonomy which has not yet been justified – nor have these

duties yet been derived. Kant hopes he can show in each case the conclusion that the maxim violates FLN, thus giving a measure of intuitive appeal to the abstract formulas he has presented. The first maxim, about suicide, violates a perfect duty to oneself. The second maxim, about making false promises to get out of difficulty, violates a perfect duty to others. The third maxim, of letting one's talents rust, violates an imperfect duty to oneself. The fourth maxim, of refusing help to those in need, violates an imperfect duty to others.

Kant's attempts to show that these four maxims violate the universalizability tests proposed in FLN have been an object of endless controversy. Some of the controversies have to do with the fact that the empirical premises Kant uses in the example are open to question; but less edifying controversies have arisen from the obviously mistaken thought that since Kant thinks the moral principle is *a priori,* he cannot be using any empirical premises at all in applying it.

Most of the controversies presuppose that Kant is proposing FUL and FLN as a wholly general test of maxims, or even as a universal decision procedure that is supposed to tell us how to act under any and all circumstances. Critics then devise maxims that are supposed to give an intuitively wrong result. Many of the resulting criticisms involve misunderstandings of FLN, of the universalizability tests, or of crucial conceptions involved in them, such as willing, willing something to be a universal law of nature, and of contradictions in volition. But other proposed counterexamples apparently do not. They show that FLN will not work as a universal moral decision procedure. Kant's self-appointed defenders, however, refuse to acknowledge this point. They seek (as if it were the Holy Grail) for some interpretation of FLN according to which all proposed counterexamples fail because they can be shown to rest on misinterpretations of the universalizability test.

Both the critics and the defenders here are wasting their time, because Kant's own application of the universalizability tests does not have the aim both sides attribute to it. His intention is only to show how certain violations of specific duties (which he makes no attempt to derive from these formulas) can be seen as cases of acting on a maxim one recognizes as opposed to what can be rationally willed as a universal law for all rational beings. The point is not to propose a universal moral decision procedure for all situations, all actions, and all maxims, but only to illustrate how some of the moral duties we already recognize can be viewed as expressing the spirit of the first and most abstract formula Kant has been able to derive from the concept of a categorical imperative. We can see how they express this spirit if we can look at some typical maxims on which people may violate recognized duties, and see how these particular maxims involve making oneself an exception to moral laws we will to be universally followed.

Kant states this point quite explicitly: "If we now attend to ourselves in any transgression of duty, we find that we do not really will that our maxim should become a universal law, since that is impossible for us, but that the opposite of our maxim should instead remain a universal law, only we take the liberty of making an exception to it for ourselves (or for just this once) to the advantage of our inclination" (G 4:424). FUL and FLN are therefore best understood in light of Kant's anthropology and philosophy of history. Their point is to oppose our unsociable propensity to self-conceit, which makes us want to see ourselves and our inclinations as privileged exceptions to laws we think all other rational beings should follow. These two formulas presuppose that we have already identified "the opposite" of our immoral maxim as such a law.

As even Kant's earliest critics were quick to perceive, FUL and FLN by themselves are inadequate to specify what these laws are. The result of dwelling on this point (as if it were something Kant needed to deny), or else of attempting to dispute it (as many Kantians misguidedly do), is only to distract attention from Kant's real aims in this discussion. Even more importantly, it draws attention away from the rest of his derivation of the supreme principle of morality in the rest of the Second Section of the *Groundwork*. For when he discusses these four examples, Kant is not finished formulating his principle. On the contrary, he has only begun. He continues his development by arriving at two other crucial thoughts, which, in addition to the concept of a categorical imperative, are really crucial to his ethical theory, namely, the worth of rational nature as end in itself and autonomy of the will as the ground of moral obligation.

Humanity as end in itself

Another side of the charge of "formalism" is the complaint that the Kantian conception of a categorical imperative is nonsensical because there could be no conceivable reason or motive for an agent to obey such a principle. Those who bring this charge have seldom even noticed that Kant's derivation of FH directly addresses this objection, by inquiring after the rational motive (*Bewegungsgrund*) for obedience to a categorical imperative (G 4:427). The first result of this inquiry is to establish that such a motive cannot be any desire or object of desire; the second result is to argue that it can only be the objective worth of rational nature regarded as an end in itself (G 4:428). Rational nature is an "end in itself" (or an "objective end") because it is an end we are rationally required to have irrespective of our desires (though Kant holds that when we have this end on rational grounds, this will produce in us various desires, such as love for rational beings, and a desire to benefit them (MS

6:401–402)). Rational nature is also an *existing* (or "self-sufficient") end, not an "end to be produced" (G 4:437). That is, it is not something we try to bring about, but something already existing, whose worth provides us with the reason for the sake of which we act. The value of rational nature is ultimate, not based on any other value. Kant thinks that the argument that something has this character can take only the form of showing us that insofar as we set ends we regard as having objective value, we already regard the rational nature that set them as having value, and we are committed to regarding the same capacity in others in the same way (G 4:428–429).

Because the worth of rational nature as end in itself is to provide a rational ground for categorical imperatives, it cannot be something whose value depends on contingencies about rational beings (such as the degree to which they exercise their rational capacities). Rather, its value must be whole and unconditional in every rational being, which entails that the worth of all rational beings is equal. Kant calls rational nature (in any possible being) 'humanity' insofar as reason is used to set ends of any kind; humanity is distinguished from 'personality', which is the rational capacity to be morally accountable. To say that "humanity" is the end in itself is to ascribe worth to all our permissible ends, whether they are enjoined by morality or not.

Kant illustrates FH using the same four examples to which he earlier tried to apply FLN. Few readers have appreciated the fact that the arguments from FH are much more straightforward and transparent than the earlier ones, and they even shed new light on the earlier arguments. Whatever objections one might raise to Kant's arguments illustrating FH, the claim that Kant's formula is empty of practical consequences is far less plausible in the case of FH than in the case of FLN. When he turns to the derivation of ethical duties in the *Metaphysics of Morals*, Kant appeals only once to anything like FUL, but well over a dozen times to FH. I submit that the reasons FUL and FLN have been treated as privileged formulations are two, both misguided. The first is simply that Kant presents these formulas first, and critical discussions have dwelt so obsessively (and inconclusively) on them that the resulting issues have served as an obstacle to considering Kant's overall argument. The second reason is the prejudice that a moral philosopher must be trying to provide us with a universal algorithm, a clever device for generating conclusions about what to do in any and all circumstances by some admirably simple process of reasoning. FH obviously cannot do that, since its application clearly depends on difficult judgments about particular cases, where it is an issue whether we are or are not treating rational nature as it ought to be treated. By contrast, FUL and FLN can be (mis)read as the sort of clever moral algorithms we were looking for. (And then we can further exercise our own cleverness – at the expense of

our comprehension of Kant's theory in the *Groundwork* – by attacking or defending the algorithms that result from these misreadings.) But let us put all this idle cleverness aside, and return to what Kant is actually doing in the Second Section of the *Groundwork*.

Autonomy and the realm of ends

Once he has derived FH, Kant can put together the thought of a categorical practical law and the thought of the rational will as a ground of value, deriving a new formula, "the idea *of the will of every rational being as a will giving universal law*" (G 4:431). Although Kant's followers as well as his critics tend to overemphasize the importance of FUL for his theory, it is hard for anyone to deny that his most revolutionary thought in moral philosophy is the idea that rational autonomy is the ground of morality. In the Second and Third Sections of the *Groundwork*, Kant himself states FA in a variety of ways, and his "universal formulations" of the moral law in the *Groundwork* (G 4:437), the *Critique of Practical Reason* (KpV 5:30), and the *Metaphysics of Morals* (MS 6:225) are all statements of FA (*not* of FUL, as they are often taken to be).[2]

As we have already noted, FUL and FLN contain only tests for the *permissibility* of individual maxims. These tests presuppose that there are universal moral laws grounding our duties, but no such law and no determinate positive duty (such as the duty never to commit suicide or positively to help others in need) can ever be derived from them. (The most their universalizability tests permit us to show is, for example, that it is impermissible to commit suicide *on this one specific maxim*.) FA, however, tells us positively that every rational will is actually the legislator of an entire system of such laws, hence that the duties prescribed by these laws are binding on us. FA says of a plurality of maxims that they collectively involve the positive volition that they (again considered collectively) *should actually be* universal laws. The universalizability tests contained in FUL and FLN provide no criterion for deciding which set of maxims, considered collectively, involves such an actual volition. (Nor does Kant ever pretend that the thought experiments involved in the four examples discussed at G 4:421–423 would ever be adequate to determine which maxims belong to this set. From Kant's procedure in the *Metaphysics of Morals*, the most reasonable surmise is that he thinks FH provides the best criterion for that.)

Kant argues that only autonomy of the rational will can be the ground of moral obligation. If anything external to the rational will were the ground of moral laws, then that would destroy their categorical character, since they could be valid for the will only conditionally on some further volition regarding this external source. (If happiness is the ground of the laws, they are conditional on our willing happiness; if the ground

of moral laws is the will of God, then their obligatoriness is conditional on our love or fear of God.)

The idea of an entire system of moral laws legislated by our will leads Kant to another idea: that of a "realm of ends" – that is, of an ideal community of all rational beings, which form a community because all their ends harmonize into an interconnected system, united and mutually supporting one another as do the organs of a living thing in their healthy functioning. FRE tells us to act according to those principles which would bring about such a system. If FH implies the equal status of all rational beings, FRE implies that morally good conduct aims at eliminating conflict and competition between them, so that each pursues only those ends that can be brought into harmony with the ends of all others.

Establishing the moral law

FA is used both in Kant's deduction of the moral law in the Third Section of the *Groundwork*, and in his alternative account in the *Critique of Practical Reason* (KpV 5:28–33). Both involve the claim that the moral law and freedom of the will reciprocally imply each other (G 4:447, KpV 5:29). This claim rests on Kant's conception of practical freedom as a causality according to self-given (hence normative) laws. To think of myself as free is to think of myself as able to act according to self-legislated principles. Kant has shown in the Second Section that if there is a categorical imperative, then it can be formulated as FA, in other words, as a normative principle self-given by my rational will. Thus if there is a moral law that is valid for me, it is so if and only if I am (in this sense) free. In the *Groundwork*, Kant argues that to regard oneself as making even theoretical judgments is to regard oneself as free, since to judge (even on theoretical matters, such as the freedom of the will) is to see oneself as following logical or epistemic norms. This means it would be self-refuting to judge that one is not free, and to represent oneself as making this judgment on the basis of good reasons. This argument is not a theoretical proof that we are free, but it does show that freedom is a necessary presupposition of any use of reason at all, and this means that any use of reason at all commits one to the validity of the principle of morality as Kant has formulated it in the Second Section of the *Groundwork*.

Notice also that this entire line of argument is wholly independent of Kant's (more controversial) idea that the causality of freedom is incompatible with natural causality, and his inference from this idea that we can presuppose ourselves to be free only by regarding ourselves as members of an unknowable noumenal world (KrV A538–558/B566–586; G 4:450–463, KpV 5:42–57, 95–106). One might agree entirely with Kant's

view that freedom and the moral law are presuppositions of reason while holding, contrary to Kant, that our freedom (in the sense of our capacity to act according to self-given rational norms) is a natural power we have that is consistent with the operation of natural causal laws.

Notice finally that Kant's conception of freedom as noumenal causality is explicitly a non-empirical conception, introduced only to solve a metaphysical problem about how the claim that we are free does not logically contradict the claim that our actions follow laws of natural causality. This conception therefore has no implications whatever for the way human moral agency is to be conceived empirically. It is misunderstood if it is treated as a metaphysical dogma about how our freedom operates. Kant's own principles rule out the possibility of our ever knowing anything about this. Kant's conception of freedom as noumenal causality is not intended to rule in or out any empirical theory about the historicity or empirical conditionedness of the development of human rational capacities or about our use of freedom in experience. If we infer from it that Kant conceives of human freedom as 'ahistorical', or not subject to variations with time and culture, then we not only draw invalid inferences from what Kant holds but we also frequently arrive at conclusions that directly contradict the actual theories of history and empirical anthropology found in Kant's own writings.

III the metaphysical system of duties

Readers of the *Groundwork* tend to emphasize FUL at the expense of Kant's later (hence better developed and more adequate) formulations of the moral law. This leads them to a picture of how Kant thinks the moral law should be applied, a picture that involves formulating maxims and ratiocinating about whether they can be thought or willed as universal laws (or, following FLN, laws of nature). When Kant finally got around to writing the *Metaphysics of Morals* (for which the *Groundwork*, as its name implies, was intended merely to lay the foundation), he provided a very different account of ordinary moral reasoning from the one suggested by this picture.

Right and ethics

The *Metaphysics of Morals (Sitten)* is divided into two main parts: the first is a Doctrine of Right (*Rechtslehre*), the second deals with "ethics" (*Ethik*), which is a Doctrine of Virtue (*Tugendlehre*). Right, which is the basis of the system of *juridical* duties, is concerned only with protecting the external freedom of individuals, and is indifferent to the incentives that lead them to follow its commands. The crucial difference between

ethics and right is that juridical duties may be coercively enforced, whereas ethical duties may not. The duties of *ethics*, concerned with the self-government of rational beings, not only require actions but also have to do with the ends people set and the incentives from which they act. They should be complied with because our reason commands us to constrain ourselves to comply with them. No authority may rightfully force us to comply with them.

Juridical duties

The basis of all juridical duties is the principle of right:

> **R**: Any action is right if it can coexist with everyone's freedom according to a universal law, or if on its maxim the freedom of choice of each can coexist with everyone's freedom in accordance with a universal law. (MS 6:230; cf. TP 8:289–290)

R bears a superficial verbal similarity to FUL, but the differences between it and all forms of the principle of morality are far more significant than the similarities. R does not directly command us what to do (or not to do). It tells us only what is *right* (*recht*) or externally just. To say that an act is "right" (i.e. externally just) is only to say that, by standards of right, it may not be coercively prevented. "Right" in this sense is not the same as the notion of 'right' used in moral philosophy (where the 'right' is distinguished from the 'good' and philosophers try to figure out which of them is based on the other). Right actions, in the present sense, include only actions which, according to the standard set up by the principle R, should not be coercively prevented, even if they are contrary to moral duty. This purely juridical standard of permissibility is not a moral standard, but is determined by what a system of right (of external justice, as coercively enforced by a legitimate authority) demands in the name of protecting external freedom according to universal law.

R no doubt *suggests* (though it does not directly state) that right, as external freedom according to universal law, is something valuable, and *implies* (though it does not assert) that we ought to confine ourselves to actions that have the property of being 'right'. If we look for Kantian reasons for these implied theses, they are not hard to find. The value that attaches to actions that are externally right is also obviously an expression of the principle of morality, as we can see most easily if we consider FH. Respect for humanity requires granting people the external freedom that is needed for a meaningful use of their capacity to set ends according to reason. That is why Kant says that the "innate right to freedom," which is the sole ground of all our rights, "belongs to every human being by virtue of his humanity" (MS 6:237). For this reason, Kant holds that

we also have an *ethical* duty to limit ourselves to actions that are right (i.e. that comply with our *juridical* duties)

Yet it is crucial to understanding R, and the notion of 'right' defined in it, to be clear that such *ethical* duties are no part of R itself, or of the juridical duties for which R serves as the principle. Both juridical and ethical duties are forms of rational self-constraint, and in this way they both fall under the heading of 'morals' (*Sitten*). But they are two distinct parts of it. Kant places 'right' ahead of 'ethics' in his exposition as if to emphasize that the two parts are distinct, and that duties of right is not merely a subclass of ethical duties, just as R cannot be derived from FA or FH or FUL, or any other formulation of the principle of morality. For juridical duties the incentive may be moral, but it may equally be prudential or (more often) something even more direct and reliable – namely, the immediate fear of what a legal authority will do to us if we violate its commands. An action fulfilling an ethical duty has greater moral merit if it is performed from duty, but the incentive from which we perform a right action makes no difference to its juridical rightness. We will have more to say about 'right', and its difference from 'ethics', in chapter 9.

Ethical duties

The *Metaphysics of Morals* conceives of ordinary moral reasoning as deliberation based on the bearing on one's action of one's various *ethical duties*. The material of one's ethical duty is constituted by "duties of virtue" or "ends that are also duties" (MS 6:382–391). In other words, for Kant, ordinary moral reasoning is fundamentally *teleological* – it is reasoning about what ends we are constrained by morality to pursue, and the priorities among these ends we are required to observe.

Thus in the *Groundwork*'s four examples, what tells us most about moral reasoning as Kant's theory presents it is not the formulation of maxims or the use of a universalizability test, but instead the taxonomy of duties through which Kant organizes the examples. The basic division is between duties toward oneself and duties toward others. Within duties toward oneself, Kant distinguishes perfect duties (those requiring specific actions or omissions, allowing for no latitude in the interests of inclination so that failure to perform them is blameworthy) from imperfect duties (where one is required to set an end, but there is latitude regarding which actions one takes toward the end, and such actions are meritorious). Perfect duties to oneself are further divided into duties toward oneself as an animal being and as a moral being (MS 6:421–442). Imperfect duties toward oneself are divided into duties to seek natural perfection (to cultivate one's powers) and duties to seek moral perfection (purity of motivation and virtue) (MS 6:444–447). Duties toward others are subdivided into duties of love (which correspond to imperfect duties)

and duties of respect (which correspond to perfect duties) (MS 6:448). Duties of love are further subdivided (MS 6:452), as are the vices of hatred opposing these duties (MS 6:458–461). Regarding duties of respect, there is a subdivision only of the vices that oppose them (MS 6:465). *Metaphysical* duties of virtue are distinguished from duties arising out of particular conditions of people or our relations to them. Kant holds that there are many important duties of the latter sort, but their detail falls outside a 'metaphysics' of morals, which deals only with the application of the supreme principle of morality to human nature in general (MS 6:468–474).

In the *Groundwork*, Kant tries (I think unsuccessfully) to relate the distinction between perfect and imperfect duties on two kinds of universalizability test involved in FLN (G 4:423–424). But he never claims that the distinction itself could be grounded on FUL, nor does he ever even try to relate FUL or FLN to the more basic distinction between duties – between duties to oneself and duties to others. Both distinctions, however, are quite easily explicated in terms of FH (cf. G 4:429–430).

A duty *d* is a duty toward (*gegen*) *S* if and only if *S* is a rational being and the requirement to comply with *d* is grounded on the requirement to respect humanity in the person of *S*. A duty is wide or imperfect (or, if toward others, a duty of love) if the action promotes a duty of virtue (an end it is a duty to set); an act is required by a strict or perfect duty (or a duty of respect to others) if the failure to perform it would amount to a failure to set this obligatory end at all, or a failure to respect humanity as an end in someone's person. An act violates a perfect duty (or duty of respect) if it sets an end contrary to one of the ends it is our duty to set, or if it shows disrespect toward humanity in someone's person (as by using the person as a mere means). Thus Kant's own moral theory (as he actually presents it in the *Metaphysics of Morals*) is much better understood in terms of FH than FUL or FLN. Corollary: attempts to construct a 'Kantian' moral theory using some interpretation of FUL as a universal test on maxims, whatever their degree of success or failure as philosophical enterprises, seriously misrepresent the actual theory Kant himself provides us.

Ends that are duties

Imperfect or wide duties should guide us in setting the ends of life. Not all ends need be duties or contrary to duty (some ends are merely permissible), but morally good people will include duties of virtue among the central ends that give their lives meaning. Kantian morality thus leaves a great deal of latitude in determining which ends to set and how much to do toward each end. The pursuit of our ends, once they have been decided upon, is constrained only by juridical duties, perfect duties to

ourselves and duties of respect to others. (In this respect, Kant's theory contrasts sharply with the terrifying rigorism of Fichte, who allows no actions to be merely permissible: every possible act is either obligatory or forbidden.[3])

In Kant's theory, the fundamental moral law is a categorical imperative, that is, a principle binding on us irrespective of any ends we may have that are independent of the principle. But as Kant interprets the fundamental moral principle, one of the main things it does is command us to set certain ends. (The ends are not presupposed by the principle as its ground, but rather they are grounded on it.) These ends, based on the categorical imperative, are exceedingly important to the structure of Kantian morality. For Kant, in fact, *all ethical duties whatever are grounded on ends.* In that sense, Kant's theory of ethical duties is entirely *teleological*, not at all *deontological* (at least if that term refers to duties that are binding on us irrespective of any end we may have set).

My own perfection and the happiness of others

There are two kinds of ends that it is our duty to have: our own perfection and the happiness of others (MS 6:385). Kant's clearest argument that we are morally required to have these ends is probably found in his discussion of the third and fourth examples he considers in the *Groundwork*, when he considers these examples in relation to FH. (FUL and FLN can never be used to show that we have any positive duties, or the duty to set any positive ends. The most it can show is that we may not adopt maxims refusing on principle to set such ends or maxims adopting contrary ends. But the imperative to treat ourselves and others as ends in themselves might require us to set certain ends regarding ourselves and others.) To treat myself as an end, I must in general honor and promote my rational capacities to set ends and develop the skills useful in furthering these ends. To treat others as ends, I must honor their rational capacities to set ends, and I do this by promoting some of the ends they set, the collective name for which is their 'happiness'.

Why do I not have a duty to promote the perfection of others and my own happiness? I have no direct duty to promote my own happiness because the concept of duty involves moral constraint, and prudential reason, quite apart from morality, constrains me to pursue my happiness. But where imprudence expresses disrespect for myself or unhappiness is likely to impair my capacity to follow principles of morality, I do have an indirect duty to promote my own happiness. What counts as the perfection of another depends on that other's choices of what ends to adopt. I cannot adopt ends for another, and have no right to constrain others to follow ends I have chosen for them. So I can have no direct duty to promote their perfection, as distinct from my duty to promote the

happiness of which their perfection, which they have adopted as an end, is a part. In other words, my duties regarding others must respect their right to choose for themselves what ends they will adopt and therefore what counts for them as their perfection. Kant's point could, therefore, be put this way: I *do* have a duty to promote my own happiness, but only insofar as my happiness falls under the heading of my perfection; and I *do* have a duty to promote the perfection of others, but only insofar as it falls under the heading of their happiness.

The general formula for ethical duties is that an action is a perfect ethical duty if omitting it means refusing to set a morally required end, or setting an end contrary to a morally required one. The analogous perfect ethical duties not to behave with contempt toward others, to defame, mock, or ridicule them, would be based on the claim that such behavior involves an end contrary to morally required ends (MS 6:463–468). Kant's theory of ethical duties is teleological, but it conceives of our pursuit of obligatory ends in a less restricted way than most consequentialist theories do. Standard devices of prudential rationality, such as summing and averaging, maximizing and satisficing, do not apply directly to our moral reasoning about the ends that ground ethical duties. My duty to promote the happiness of others is *not* a duty to *maximize* the collective happiness of others. It leaves me with quite a bit of latitude to decide whose happiness to promote, and which parts of their happiness to promote. My duty to promote my own perfection is not a duty to achieve any specific level of overall perfection, much less a duty to make myself as perfect as I can possibly be. Kant's theory leaves it up to me to decide which talents to develop and how far to develop them. Kant's theory gives us no reason even to reproach a person for being less virtuous or morally perfect than they might have been.

All duties of virtue are, in their concept, wide, imperfect, and meritorious duties (MS 6:390–391). I behave meritoriously insofar as I act to promote an end falling under the concept of the required ends. But I deserve no blame for failing to promote the end on any given occasion, and *a forteriori* no blame for not promoting it maximally. In general, it is up to me to decide whose happiness to promote, and to what degree. Ethics allows me latitude or "play-room" (*Spielraum*) in deciding such matters (MS 6:390). Thus moral agents themselves, as free agents, and not the theory of moral principles or duties, are responsible for the design of their individual life plans.

Because the ends morality requires us to adopt are general *kinds* of ends and not specific ends, and because the requirement is to set ends of those kinds rather than to maximize any kind of good, a Kantian theory of duties does not threaten to be inhumanly demanding on us, as consequentialist or utilitarian theories of moral duty threaten to be. This point has seldom been appreciated, probably because attention has been

distracted from it by some of Kant's infamously extreme opinions about certain duties, such as the duty not to lie. But it is very questionable whether Kant's convictions about specific topics really follow from his ethical theory. In my opinion, Kant's theory, if it is correctly understood, seems more vulnerable to the charge that it is too lax than to the charge that it is too strict. The chief means Kant has for rebutting the charge is to appeal to specific contexts of action, or to specific institutional relationships in which we stand to others, to render our duties to them stricter and more precise. Kant's chief idealist followers, Fichte and Hegel, correctly took this route, by relating ethical duties to a rational social order and to the roles individuals are supposed to play in it.

Ethics as virtue

The title of Kant's system of ethical duties is the "Doctrine of Virtue." His name for the obligatory ends of pure practical reason is "duties of virtue." In the *Critique of Practical Reason*, Kant describes "virtue" as "a naturally acquired faculty of a non-holy will" (KpV 5:33), or, more specifically, as "the moral disposition in the struggle" (*im Kampfe*) (KpV 5:84). In the *Metaphysics of Morals*, virtue is characterized as "the moral strength of a human being's will in fulfilling his duty" (MS 6:405; cf. 6:394). "Moral strength" is an "aptitude" (*Fertigkeit, habitus*) in acting and "a subjective perfection of the power of choice" (MS 6:407). Obligatory ends are called "duties of virtue" because virtue is required to adopt and pursue them. There is only a single fundamental *disposition* of virtue, but because the ends which it is our duty to have are many, there are many different virtues (MS 6:383, 410). I can have one virtue and lack another if my commitment is strong to one obligatory end but weak to another.

Kant holds that we have a duty to cultivate feelings and inclinations that harmonize with duty and to acquire a temperament suitable to morality (MS 6:457). But he does not equate *virtue* with success in fulfilling that duty (MS 6:409). Virtue is needed precisely to the extent that good conduct is hard for us, since it consists in the strength we need to perform a difficult task. A person might have a temperament so happily constituted that their feelings and desires make duty easy and pleasant to do. Such a temperament is not virtue, but only makes virtue less often necessary. The person may still be virtuous too, but virtue is a quality of *character* (of the active strength of rational maxims), not of temperament (of the feelings and desires we passively experience).

This conception of virtue follows naturally from Kant's theory of human nature. For according to this theory, in society our inclinations, as expressions of competitive self-conceit, are inevitably a counterweight to the moral law, which requires strength to overcome it.

Therefore, there can be no reliable fulfillment of duty without (some degree of) virtue. The theory of ethical duties is called a 'Doctrine of Virtue' only because human nature is such that virtue is the fundamental presupposition of all reliable ethical conduct. In the civilized condition, where our feelings and desires are corrupted by social competition and self-conceit, it would be not only dangerous, but blamably irresponsible, to rely (as Hutcheson and Hume would have us do) solely on non-rational feelings and empirical desires as the motives for morally good conduct.

further reading

Marcia W. Baron, *Kantian Ethics (Almost) Without Apology*. Ithaca: Cornell University Press, 1995.

David Cummiskey, *Kantian Consequentialism*. New York: Oxford University Press, 1996.

Paul Guyer, *Kant on Freedom, Law and Happiness*. New York: Cambridge University Press, 2000.

— (ed.), *Kant's Groundwork of the Metaphysics of Morals: Critical Essays*. Totowa, NJ: Rowman and Littlefield, 1998.

Barbara Herman, *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press, 1993.

Thomas Hill, Jr., *Dignity and Practical Reason*. Ithaca: Cornell University Press, 1992.

Christine M. Korsgaard, *Creating the Kingdom of Ends*. New York: Cambridge University Press, 1996.

Robert B. Louden, *Kant's Impure Ethics*. New York: Oxford University Press, 2000.

Allen W. Wood, *Kant's Ethical Thought*. New York: Cambridge University Press, 1999.

— (ed. and tr.), Kant, *Groundwork for the Metaphysics of Morals*, with essays by J. B. Schneewind, Marcia Baron, Shelly Kagan, and Allen Wood. New Haven: Yale University Press, 2001.

notes

1 However, it remains puzzling to me how Kant could have said what he did in his infamous essay on the right to lie, given some other things he says about the duty to tell the truth, when it applies and when it doesn't. See chapter 9, note 2. Few who express condescension or horror at Kant's famously inflexible views about lying even seem to be aware of this puzzle.

2 With regard to the *Critique of Practical Reason*, this point has been noted by both H. J. Paton, *The Categorical Imperative* (New York: Harper & Row, 1949), p. 130, and Lewis White Beck, *Commentary on Kant's Critique of Practical Reason* (Chicago: University of Chicago Press, 1960), p. 122 and note 22.

3 See Fichte, *System of Ethics, Fichtes Sämmtliche Werke*, ed. I. H. Fichte (Berlin: Walter de Gruyter, 1970), 4:156, 204, 264.

chapter 8

the theory of taste

I why a third 'critique'?

The *Critique of Pure Reason* was foundational for the philosophical contributions that make Kant's work memorable for us. The *Critique of Practical Reason* was an outgrowth of Kant's work on a second edition of this foundational work, and also of his attempts to clarify the foundations of practical philosophy as he had presented them in the *Groundwork*. It is harder to say why Kant wrote the *Critique of the Power of Judgment*, which, he says, brings his entire critical enterprise to an end (KU 5:170). His fundamental and clearly avowed purpose was to bridge what he perceived to be a yawning gulf between the treatments of theoretical and practical reason in his philosophy, and thereby to unify his philosophical system. But exactly what his solution to this problem is supposed to be, or even what the problem itself is supposed to be, are matters of deep dispute among Kant scholars down to the present day. In a study of this kind, I will avoid stating any opinion on these questions. For any account I might give would be unavoidably controversial, and there would be no space here to explain or defend it. (Perhaps I can be certain of drawing criticism from all sides, however, if I offer the modest suggestion that from the standpoint of Kant's enduring philosophical legacy, the dark issues surrounding the unity of the Kantian system may all be of less interest than devoted Kant scholars usually take them to be.)

Apart from this fundamental (if obscure) purpose, however, Kant's aim in his third and final *Critique* was also to address two topics that were of great philosophical importance in his time, and to rein in some of the things that were said about them that he thought violated the critical strictures his philosophy had put in place. The first such topic was taste, its proper standards, and the implications of our experience of beauty for metaphysics and morality. These were subjects with which much eighteenth-century thought had been creatively occupied. The second topic was natural teleology, its function in natural science, and its implications for both morality and religious belief. The mechanistic

view of nature championed by Descartes and much of early modern scientific thought had been countered by Leibniz and the Cambridge Platonists, and the second half of the eighteenth century witnessed a strong continuation of this reaction, especially among certain German thinkers. Kant wanted to give its due to this reaction, while curbing some of the anti-scientific enthusiasm to which he thought it was prone.

But both topics, as Kant could clearly see, were pertinent to problems within his own critical philosophy itself that he regarded as still outstanding. Kant named these problems collectively by referring to the "incalculable gulf" between the sensible and the supersensible, nature and freedom, theoretical and practical reason (KU 5:175–176). Almost immediately upon the reception of Kant's critical philosophy, and ever since, Kant has been charged by some with establishing a set of false and unhealthy "dualisms" – between appearances and things in themselves, nature and morality, inclination and duty. The *Critique of the Power of Judgment* is Kant's own acknowledgment of these criticisms, and his attempt to answer them.

As with most of Enlightenment aesthetics, Kant holds that there is a close connection between morality and aesthetic feelings for the beautiful and sublime. He sees these feelings as connecting and mediating between moral reason and our sensitive nature. Beauty and sublimity give us an authentic feeling for morality, and even (in Paul Guyer's felicitous phrase) an *experience of freedom*.[1] As we shall presently see, in the experience of beauty, on Kant's account of it, the faculty of judgment mediates – and registers a spontaneous harmony between – our sensible faculty of imagination and our intellectual faculty of understanding. The experience of beauty also affords us awareness of "aesthetic ideas" – imaginative or sensory representations to which no concept is adequate, which are the complement to ideas of reason – intellectual representations to which no sensory intuition can be adequate. The other main theme of the third *Critique*, natural teleology, ultimately connects our theoretical science of nature to the system of moral ends, opening up for us a vision of nature that harmonizes with our moral vocation, thus bridging the gulf between theoretical and practical reason, freedom and nature.

At the same time, however, the *Critique of the Power of Judgment* carefully defines the unique place of aesthetic judgment in our mental life, and determines the special, and limited, role of teleological judgment in our inquiry into nature. In this way, it also carries out Kant's *critical* enterprise by guarding the boundaries of our cognition against a kind of aesthetic or religious enthusiasm that he regards as both theoretically irresponsible and practically dangerous.

In this chapter, we will briefly survey the aesthetic side of Kant's project in the third *Critique* and its relation to the traditional theories of

taste between which Kant was trying to mediate, and whose opposition he was also trying to transcend.

II judgments of taste

Judgments that a certain object is beautiful or ugly are very peculiar. They are not like mere assertions that you find something likable or repugnant – what Kant calls judgments about the 'agreeable' or 'disagreeable'. Suppose I like the taste of chamomile tea because I associate it with my sweet old grandmother who used to serve it to me as a child with marzipan, while you dislike it because your prune-faced governess used to force it on you as a sort of punitive medication accompanied by a scolding whenever you did something she considered bad for your health. There is nothing here for us to disagree about. I like it and you don't. We both acknowledge these facts, and know why they obtain, and that's all there is to it. Yet if I say that a certain object is beautiful and you say it is ugly, we do take ourselves to be disagreeing about something – we both think that if one of our assertions is true, then the other has to be false. Both assertions seem to be predicating of the object a real and objective property, a property like its size, or its mass, or even whether it is good or bad. All these judgments Kant, at any rate, understands as predicating of the subject objective properties of things. A one-pound sack of rice is objectively one pound. A knife that effectively serves the standard purposes of knives is instrumentally or functionally good, while one that makes those purposes harder to attain is bad; an action forbidden by moral laws is morally bad, while one they determine to be meritorious is morally good.

But judgments of beauty and ugliness are not objective judgments, in the way that judgments about size or weight are, or even in the way that judgments about goodness and badness are. The essential thing about a beautiful object is that it pleases us, and about an ugly one that it displeases us (when the objects are considered in the appropriate way and our experience of them is not distorted by factors alien to a genuine judgment of taste). Our judgments of taste *disagree* because we each think that the other ought at least to be able to consider the object in this appropriate way, and we think that when this happens, the other ought then to be pleased by the object we consider beautiful and displeased by the object we consider ugly. Pleasure and displeasure, however, are essentially subjective feelings. As Kant often insists, the mere fact that an object pleases or displeases gives me no indication whatever of its objective properties, not even in the way that the color of an object tells me something objective about the light its surface reflects under certain conditions. The only information it gives is about the subject.

Judgments of taste thus seem to present us with a paradox. They can't possibly be objective judgments about an object, but we treat them as though they were. Beauty and ugliness function in discourse as if they were objective properties of things, even though we know perfectly well that they can't be. How does it happen that we come to consider being pleased or displeased by an object, at least when it is regarded in a certain way, as if it were an objective property of the object? It would seem that we treat this special sort of pleasure or displeasure as normative, as though when so regarded it were in some sense *right* to be pleased by a beautiful object and *wrong* not to be pleased by it.

Kant is often viewed as trying to mediate or transcend the opposition between 'rationalism' and 'empiricism' in the theory of knowledge. But the only area of philosophy in which he understands his own task in precisely these terms is aesthetics (KU 5:346). It may help us to understand his theory of taste if we look briefly at the 'rationalist' and 'empiricist' solutions to the paradox about judgments of taste, and see why Kant found both solutions unsatisfactory. Rationalists such as Baumgarten or Mendelssohn identify beauty with goodness or perfection as apprehended through the senses rather than the intellect. The distinctiveness of this sensory mode of apprehension for them lies not only in the fact that it is confused rather than distinct, but also in the fact that they regard sensory pleasure and displeasure as indispensable to motivating us to action. Empiricists such as Hutcheson or Hume identify the beautiful with the agreeable, but agreeableness felt only under certain idealized conditions – free from interest or bias, by someone experienced in the kind of aesthetic object that is to be judged. Kant rejects the rationalist account because it locates the distinctively subjective and non-conceptual character of beauty only in the mode of its apprehension, whereas these features belong to the nature of beauty itself. He regards the empiricist account, by contrast, as unable adequately to account for the normativity of aesthetic judgments, since the canonical status of the idealized conditions of aesthetic judgment count as normative only because judgments of agreeableness made under these conditions happen, once again, to elicit our approval – because the conditions themselves are agreeable to our notions of aesthetic judging – but this too can never be more than yet another empirical fact about what we find agreeable.

A satisfactory solution to the problem of taste must combine the genuine normativity, or universal validity and necessity, of such judgments for all subjects, with their essential subjectivity – with the fact that they always refer fundamentally to what pleases or displeases, and not to any objective property that occasions pleasure or displeasure merely because of the cognitive faculty through which it is cognized.

Kant's proposed solution consists in exploiting his own account of our cognitive faculties. As we saw in chapter 2, human cognition occurs

through the co-operation of sensible intuition. Through it objects are given, and through the understanding they are conceptualized, making objective judgments about them possible. In considering our capacity for aesthetic judgments, Kant is concerned with the understanding's relation not to intuition but to imagination, since aesthetic judgments do not concern the *existence* of objects (as given in intuition) but only their sensible *representation* for imagination, whether they are given as existing or not.[2] Judgment is the faculty that relates what is given in imagination to the concepts under which it may be brought. This relation may be of two kinds. In *determining* judgment, a concept is applied to what is given, while in *reflecting* judgment, a concept is sought out for what is given. In both cases, what is required is some sort of match or harmony between representations of intuition or imagination and conceptual representations, which rests in turn on a harmonious operation between the faculties of imagination and understanding themselves.

Presupposed, therefore, by any act of judgment is the operation, in relation to each other, of both imagination and understanding on a given representation. Especially in reflecting judgment, this operation does not presuppose any given concept, since the point of reflecting judgment is to arrive at one. But judgment also involves what Kant calls the "free play" of imagination and understanding in mutual relation – *free*, that is, from guidance by any concept. Some representations are such that already in this free play, they bring imagination and understanding into harmonious relation. What the imagination represents is then spontaneously well suited to the understanding's conceptualization of it – and the subject experiences this spontaneous harmony between the imaginative and intellectual faculties even prior to the application of any given concept. The experience of this harmony animates, quickens, or enlivens (*belebt*) both faculties, because the life activity or functioning of each is most successful, in that sense most animated or alive, when they work in harmony. And for any of our faculties, the subject experiences its successful exercise in the form of the feeling of *pleasure*. Hence the harmony or mutual animation of imagination and understanding in free play takes the form of a pleasure. This feeling of pleasure, in Kant's theory, is aesthetic pleasure, or the experience of *beauty*. The contrary aesthetic judgment, one that finds something to be ugly, is in place when the representation impedes the understanding in its grasp of what is given in imagination, so that our faculties cannot smoothly co-operate in the cognition of what is given to them even prior to any concept, even if there are suitable concepts ready to hand for the ugly object.

The judgment of taste that finds something beautiful or ugly must be subjective and yet have universal validity. Kant's theory explains this feature through the fact that beauty is experienced in a representation that produces pleasure solely by animating our faculties in free play,

free from determination by any concept. Since our faculties are in play free from any concept, the source of the aesthetic pleasure is purely subjective – it is independent of any objective *judgment* about an object, which would depend on a determinate concept serving as the predicate. Yet because it depends solely on the essential nature of imagination and understanding in general – which are the same in all human beings – the conditions under which this pleasure is felt will be universally valid for all who experience the representation in terms of the free play of their faculties.

Aesthetic pleasure is therefore also *universally communicable* – it is a pleasure we can share, and expect to share, with all subjects who entertain the same representations in a purely aesthetic manner, that is, through the free play of their imaginations and understandings in mutually animating harmony. And because we are also sociable beings, who take pleasure in being able to communicate with others of our kind, there is added to our pleasure in the mutual animation of our faculties also a pleasure of sociability – the pleasure of having sensations that are universally communicable, sharable with others (KU 5:216–219).

In a crucial section of the Analytic of the Beautiful, Kant asks whether in the judgment of taste the feeling of pleasure precedes or follows the judging of the object as beautiful (or ugly) (KU 5:217). His (perhaps surprising) answer is that the judging has to precede the pleasure, since otherwise the pleasure could be only agreeableness and not pure aesthetic pleasure.

This implies two other conclusions about aesthetic pleasure that are noteworthy. First, aesthetic pleasure requires a certain reflectiveness. It is part of aesthetic pleasure itself that we are conscious of it as possessing a kind of universal validity. If our aesthetic judgment that something is beautiful is a genuine aesthetic judgment, and a correct one, then we are aware that any other subject who makes a genuine aesthetic judgment about the object must also take aesthetic pleasure in it and judge it to be beautiful. Moreover, this awareness is itself an ingredient in the aesthetic pleasure itself, not a mere optional addition to it. Second, the experience of aesthetic pleasure always has quite direct reference to its communicability to others, and therefore to our sociability. It is essential to our enjoyment of the beautiful that we are aware of it as something others can (and even should) enjoy as we do. It follows that the enjoyment of the beautiful thus cultivates or educates us, and this in two distinguishable but related ways: on the one hand, it promotes our cognitive powers – in particular, the harmony between our imagination and understanding; on the other hand, it also cultivates our mental powers for sociable communication. And Kant's account of aesthetic pleasure shows these two forms of mental cultivation to be closely related to each other. In other words, it is a deep part of Kant's conception

of our cognitive capacities that sense and understanding are meant to work together in harmony, and that their exercise must be social in both context and destination. This point of Kantian doctrine gives the lie to those who accept the image of Kantian philosophy as built on harsh dualisms between sensibility and understanding, and as individualistic in its conception of human knowledge and agency.

Representations of imagination are beautiful when they involve shapes, regularities, symmetries, contrasts that appeal to our understanding even apart from conformity to any concept under which we might want to bring them. This is especially true when these patterns involve a development through time, as in the case of the successive notes in the melody or chord changes in a piece of music, the succession of words in a poem, or the composition of a painting when we survey its parts successively and then come to grasp it as a whole. These features of a beautiful object constitute what Kant calls a "form of purposiveness" (KU 5:221) – that is, when they are so related to one another that their relations involve a coherence that might have been the product of a design readily grasped by the understanding. But this must be – in Kant's other paradoxical phrase – a "purposiveness without a purpose (or end)" (KU 5:220), since the purposive relations or coherence is grasped apart from any concept specifying a determinate end.

Kant aims to capture the elements of truth in both the rationalist and empiricist theories of taste while correcting the shortcomings of each and going beyond both. The rationalist locates beauty in sensibly apprehended *perfection*. Perfection is certainly closely akin to the formal purposiveness of beauty on Kant's account, since the perfection of a manifold consists in its agreement or unity (KU 5:227), and this is what leads the understanding to experience a manifold of imagination as beautiful. Beauty, therefore, is analogous to the perfection of a living thing (KU 5:375). The universal validity of the aesthetic judgment is similar to the validity of an objective judgment telling us that a particular thing conforms to the excellence of its kind denoted by the concept under which we bring it. But properly speaking, perfection always presupposes such a concept of the unity constituting that kind of thing, and indicates the conformity of the thing to this concept. This determinate concept, however, is missing in the case of pure judgments of taste.

Kant's account shares with the empiricist account the idea that what is beautiful must please *subjectively* under conditions appropriate to the properly aesthetic judging of it. But it identifies these conditions not merely with those under which we might happen contingently to approve of the judgment as properly aesthetic, but with those under which our pleasure or displeasure is occasioned solely by the free play of imagination and understanding, which provides the universal validity of a judgment of taste.

Kant's theory of pure judgments of taste has usually been understood as favoring an austere, formalistic aesthetics, in which the abstract form of beautiful objects – in particular, of works of art – predominates in aesthetic importance over content, and excludes such pleasing features of aesthetic objects as charm and emotion. It is true that Kant specifically distinguishes the effect of beauty from both charm and emotion (KU 5:223), and regarding pure judgments of taste, he attributes the beauty of objects to their "form of purposiveness" (yet without a determinate end) rather than to anything about their content or the human ends they may serve (KU 5:221–222). But Kant's theory by no means ignores the content of works of art, and in dealing with such works it has resources enabling it to take account of the aesthetic significance we attribute to content, functionality, and the emotional evocativeness of artworks.

Kant was aware of Hume's thesis that the beauty of artworks especially is related to their utility, and while he thought the thesis is false regarding pure judgments of taste, he wanted to grant its truth within the appropriate sphere. For this reason, he distinguishes the "free beauty" ascribed by a pure aesthetic judgment from "adherent beauty" ascribed to something on the ground that it conforms excellently to the concept of its kind (KU 5:229). Thus a beautiful horse or summer house may be judged in accordance with the appearance that is taken to conform to the species of excellence indicated by the concepts of these kinds of things, even by their utility in human life. But pure aesthetic judgments are those that do not presuppose any concept.

A flower is beautiful merely because of the shapes and colors apprehended when we look at it. If subsequently we also form a concept that is to be normative for the beauty of, say, a tulip or a begonia, this involves the superimposition of a standard of adherent beauty upon the free beauty belonging to the flower simply as a beautiful product of nature. The same is true regarding works of human art that have a function – for example, a beautifully designed house or eating utensil. The free beauty of its form, as the object of a pure judgment of taste, is distinct from aesthetic judgments about it as conforming to a concept of its kind or relating to its function. It is noteworthy that the same is true regarding works of art that have a moral purpose – for instance, the eloquence of a sermon or moral exhortation, whose serviceability to awakening noble sentiments in its hearers pertains only to its adherent beauty, and is distinct from the free beauty that might belong to the formal properties of the way it uses language, or of the images and metaphors it employs.

In light of this, the claim that Kant privileges formal properties of aesthetic objects over their content or their capacity to arouse emotions would have to depend on the thesis that pure judgments of taste regarding free beauties are to be privileged over judgments of taste relating to

adherent beauties. It is probably true that Kant accepts this thesis, but our grounds for thinking so are at best indirect, having to do with the emphasis he places within his theory on pure judgments of taste, and the relatively small amount of attention he devotes to adherent beauty. If we look in Kant's texts for any explicit assertion of the thesis, and especially for any argument for it, I think we will not find any.

Another controversial part of Kant's aesthetic theory – a point raised early on by his erstwhile student Herder – is his claim that aesthetic judgments may pretend to universal validity. Some have thought this flies in the face of the cultural relativity of aesthetic values, and also of the fact that people are often at least as much interested in forming their own unique personal tastes as they are in claiming that what they experience as beautiful ought to be so experienced by others. But Kant has some very cogent replies to these objections once we realize that any individual's capacity for aesthetic judgment – especially in the case of humanly created art works – is bound to be limited by the individual's past experience, cultural conditioning, and the consequent extent of their acquired ability to make the pertinent pure judgments of taste. People brought up in different traditions of musical appreciation may lack the ability to judge musical works in a tradition foreign to them. Yet this does not mean that the pure judgments of taste made by someone with the requisite background would not be valid for them, and would agree with the judgments they would make if they were to acquire the proper connoisseurship in the alien tradition.

Likewise, if we are mature and cultivated judges, our aim in cultivating our own taste is not to assert our own idiosyncrasies or display our differences from others, but rather to develop our particular aesthetic expertise in ways that take account of the inevitable limitedness of our background and perspective. People also develop their own expertise in theoretical matters, some specializing in mathematics or biochemistry, others in medieval Latin paleography, still others in seventeenth-century English political history. The truths each discovers (to the extent that they find the truth about these subjects) are valid for all others (otherwise they would not be truths), even for others who lack either the desire, the training, or even the native talent to come to know them. In the same way, I may focus on developing my taste for European classical music rather than jazz or Indian classical music without being in any position to declare these other musical traditions to be without aesthetic validity. And if I did presume to declare this without possessing the requisite expertise to judge about them, my declaration need not be taken any more seriously than that of a historian who might foolishly assert that there is no truth at all in some branch of mathematics she has never studied. Questions of taste arising among people of the same culture and experience judging the same or similar objects far more often

have determinate answers that seldom raise questions of relativity or incommensurability: Mozart's music is superior to Salieri's; and anyone who prefers Lawrence Welk to Duke Ellington ought to be ashamed to admit it.

III beauty and morality

Kant's theory of taste tries to unite the elements of truth in both the rationalist and empiricist theories. Yet it also differs from both the rationalist and empiricist theories in one crucial way. Both theories in the end identify beauty with moral goodness, since for the rationalist, goodness consists in perfection, while for moral sense theorists such as Hutcheson and Hume, goodness is simply identified with what excites our disinterested approval. Hutcheson, at any rate, even regards moral and aesthetic approval as operations of essentially the same sentiment. (Hume's theory of moral approbation, which involves a distinction between natural and artificial virtues and the role of judgments of utility in producing moral sentiment, is more complex.) For Kant, however, the judgment of taste is sharply distinguished both from subjective agreeableness and from all objective judgments about goodness – whether instrumental or moral. From one point of view, this means that Kant's theory of taste secures what some have called (using Kantian terminology, but in a way Kant himself never does) "the autonomy of the aesthetic." By this is meant that aesthetic judgments are treated as having their own standards. Standards of beauty or aesthetic merit are not only distinct from but also independent of all the standards of morality or utility or any other sort of *goodness*. In this way, Kant is seen as leading out of eighteenth-century aesthetics, which generally viewed beauty generally and art in particular in terms of their functions in moral psychology and moral education, and into a new, more modern and liberated aesthetics in which art is seen as having its own independent function in human life apart from morality or any other goodness-oriented enterprise.

Viewing Kant's theory of taste in this way may accurately indicate its role in inspiring a subsequent tradition in art and aesthetics. But it utterly misses the point as far as Kant's own view of the matter is concerned. For Kant belongs firmly within the tradition of eighteenth-century aesthetics in thinking that the real significance of beauty and taste for human life is chiefly a moral significance. The importance for Kant of the so-called "autonomy of the aesthetic" is that it is only when judgments of taste can be distinguished from moral judgments that they can be understood as playing the distinctive and positive role that they do play, and should play, in the moral life.

Kant understands *interest* as a pleasure taken in the existence of an object (or state of affairs). Aesthetic pleasure is disinterested, because it is a pleasure in the mere representation of an object, irrespective of its existence (KU 5:204–205). For instance, our purely aesthetic pleasure in the architectural design of a house is pleasure that we take whether or not we expect to live in it, or even expect it to be built at all. Moral motivation in an action or a morally good end is also disinterested, in the sense that it is not based on any subjective agreeableness for us of the existence of the action or the end. But when we perform an action or pursue an end because it is morally good, our pleasure in it is involved with an interest, since our awareness of its goodness gives rise to a desire that the action should be performed or the end achieved, and this pleasure does relate to the hoped-for existence of an object. Aesthetic pleasure arises neither from the agreeableness nor from the moral goodness of the existence of anything. In that way, the pleasure we take in beauty is distinct from the pleasure we take in moral goodness, and aesthetic value for Kant is (as some say) "autonomous" in relation to moral value.

Yet although aesthetic pleasure is disinterested in this sense, Kant is very clear that it can and does give rise to interests, even to interests closely connected to moral interests. Our sociability, and our moral interest in reaching harmony or agreement with others, gives us what Kant calls an "empirical interest in the beautiful" (KU 5:296–298). That is, both sociability and morality lead us to place special value on feelings that can be shared with and communicated to others, and that we recognize as universally valid for all subjects. Our valuing of such feelings over private feelings of mere agreeableness is educative; it teaches us to value, even on the level of feeling, what conforms to universal standards, just as on the level of cognition we ought to value an assent that can be justified by reasons valid for others over mere private persuasion based on prejudice or self-interest, and on the level of morality we should value principles that are valid for all rational beings over mere maxims that lack the form of universal law. Kant thinks we also take an "intellectual interest in the beautiful" (at least in the beautiful in nature). That is, our experience of beautiful objects in nature (such as a flower, a bird, or a butterfly) *creates an interest*, closely allied to our interest in the morally good, in the *existence* of such objects (KU 5:298–300). Aesthetic judgments are thus closely related to what Kant sees as our moral duty to value and promote what is beautiful in nature (MS 6:443). Thus our aesthetic pleasure is itself neither interested nor grounded on any interest (even a moral one), but natural beauty *produces* in us an interest, which, Kant says, is "always the mark of a good soul" (KU 5:299).

One of Kant's most distinctive (also somewhat puzzling) doctrines in this area is that beauty itself is a symbol of morality (of the morally good)

(KU 5:351–354). For Kant, a 'symbol' is a way of giving intuitive content to an *a priori* concept. When such content is specified in such a way as to enable us to recognize an instance of the concept, it is called a 'schema'. But when the concept is an idea of reason, to which no intuition could ever be adequate, intuitive content cannot be given to it directly but only by analogy, and the representation through which this is done is called a 'symbol'. Kant's concept of a symbol is closely related, therefore, to his adoption of the scholastic theory of analogical predication (to which Kant subscribes in discussing our application of empirical predicates to God). Morality or the morally good, however, is also an idea to which no sensible intuition can ever be adequate, so symbolism is in place here too. A symbol is a predicate applied to something not because the predicate is literally true of it, or even because it resembles what is literally true of it, but rather because the procedure of the understanding in thinking this intuitive predicate bears some analogy to its procedure in thinking the idea of reason.

Kant indicates four such ways in which beauty can symbolize morality. First, the beautiful pleases immediately, just as the morally good is valued immediately for its own sake (though based on a concept, not on a harmony of the faculties in free play). Second, the beautiful pleases disinterestedly, just as the morally good appeals to us apart from any antecedent interest (though recognition of something as good involves taking an interest in it). Third, enjoyment of the beautiful involves a freedom of the imagination that is nevertheless in conformity with understanding, which is analogous to what occurs in a morally good action, where our faculty of desire is in free conformity with laws of reason. Finally, pleasure in the beautiful is universally valid, just as the principle of a morally good action conforms to universal laws (KU 5:354). Enjoyment of the beautiful is therefore an experience that is capable of reminding us of what is morally good, and representing morality in a way that appeals to our feelings and imagination. For Kant, the whole point of the independence or 'autonomy' of the aesthetic in relation to morality is that aesthetic pleasure adds a new dimension to our moral experience, something it could not do if it were nothing but a sensible apprehension of perfection or an exercise of the same sentiment of approval that grounds morality.

IV the sublime

Eighteenth-century aestheticians often distinguished between two contrasting forms of aesthetic experience – the beautiful, which pleases through some sort of perfection, harmony, or purposiveness, and the sublime, which pleases despite (or even because of) the way it

exceeds our capacity to comprehend it and even threatens to overwhelm us with its might. Interest in the topic of the sublime arose in the modern period with the rediscovery of the ancient treatise *Peri hypsous*, attributed to the grammarian Longinus, and its translation by the French aesthetician Nicolas Boileau. The object of Longinus' discussion was a certain kind of elevated rhetorical style, and its intended effect on the mind. But modern discussions soon turned their attention to the larger experience of sublimity not only in rhetoric but also in art and in nature. This is the experience of the emotions of awe and wonder that can be awakened by anything perceived as great and terrifying. To eighteenth-century aestheticians, the experience of the sublime seemed obviously distinct from that of the beautiful, and yet also to be an important experience, involved both in art and in our aesthetic experience of nature, that needs to be understood alongside the experience of the beautiful.

Kant distinguishes two forms of the sublime: the *mathematically* sublime, which we represent as "absolutely great" (the wide ocean, the majestic vault of heaven, the thrill of looking up at a colossal overhanging cliff) and the *dynamically* sublime, which we represent as absolutely powerful (the ocean storm, thunder clouds, volcanoes). Kant appears to think that we make judgments about the sublime in nature that purport to have a universal subjective validity (KU 5:248), but the focus of his discussion is not on how we judge *which* natural objects are sublime. The puzzle is rather how it happens that we can take *pleasure* in the sublime, since its chief effect on us seems to be to frustrate our understanding (by exceeding its capacity to comprehend) or our will (by threatening to overpower us). The feeling of the sublime has a strange capacity to *move* us – Kant says that we say of those who are not pleased by the beautiful that they lack *taste*, but of those who remain unmoved by the sublime that they lack *feeling* (KU 5:265). The most interesting question for Kant is: What is the significance for human nature of the fact that we are moved by the sublime?

Edmund Burke's psychological explanation of this phenomenon, which inspired much of the eighteenth-century discussion of the sublime, appeals to the fact that we experience the sublime only when we are in fact safe from it. A person who is about to drown in a raging tempest feels fear, not sublimity (or he would be able to experience the sublimity of his situation only insofar as he could cease directly to fear it). Burke thinks our pleasure in the sublime lies in the contrast between its fearfulness and the actual position of safety that we occupy when we experience the sublime.

Kant is unsatisfied with such a merely psychological explanation, since it fails to account for the normativity or quasi-objectivity in the experience of sublimity – the fact that the feeling of the sublime has a certain importance for us, such that a person who could not experience it

would seem to us to be lacking in some important respect. Kant locates the importance of the sublime in the fact that it is a felt awareness of our moral nature. Moral ideas transcend every capacity of sensibility to represent what might be adequate to them. The feeling of the mathematically sublime is our way of experiencing this transcendence (KU 5:257–258). Although we are finite beings of nature, our moral vocation gives us a worth infinitely greater than any that could be drawn from mere nature. The feeling of the dynamically sublime is our way of experiencing the infinite superiority of our supersensible moral disposition to any power that nature might exercise over our bodies. "Thus sublimity is not contained in anything in nature, but only in our mind, insofar as we become conscious of being superior to nature within us and thus also to nature outside us" (KU 5: 264).

It is a natural thought, also a common enough one, that the feeling of the sublime is closely allied to religious experience, especially to human awe and fear in the presence of God's majesty, almightiness, and punitive justice – or as Rudolf Otto was later to characterize it, the experience of the 'numinous'. It is therefore striking that Kant is conspicuously cool toward any such associations. Kant thinks that those who associate the feeling of the sublime with God are likely to exemplify the popular but contemptible religious disposition that seeks divine favor through degrading attempts to ingratiate oneself like a wheedling sycophant groveling before a pompous cosmic tyrant – a religious disposition, in other words, that dishonors both God and ourselves (KU 5:264). Kant thinks that a self-respecting human being has no reason to be afraid of God. The right attitude toward God insofar as we are aware of our moral imperfections is not pious terror or cringing repentance but instead a sober resolve, grounded in our awareness of our moral freedom, to do better in the future. What we experience as transcending the power of nature is not the 'numinousness' of a powerful alien Being who takes some sort of sadistic pleasure in overwhelming and terrifying us, but rather the sublimity of our own moral freedom. The truly sublime object to which our aesthetic experience relates is therefore not God but our own moral disposition and vocation.

V art and genius

Modern aesthetics has usually thought of itself as identical to the philosophy of art. But in considering Kant's aesthetic theory, it is crucial to realize that for him the object of the most significant aesthetic experiences is not human artworks but *nature*, whether beautiful nature or sublime nature. Nearly all of Kant's own examples of beauty or sublimity are drawn from nature, not from art. The fact that

Kant's account is oriented more toward natural than artistic beauty helps to explain, for instance, his apparent underemphasis on cases that might be thought to raise issues about the incommensurability or diversity of taste; for such cases arise much less often with natural beauty than artistic beauty.

Although he recognizes that artistic beauty is closely allied to our sociability, just as the cultivation of taste is allied to the cultivation of our capacity for universal communication, Kant thinks of the appeal made to us by beauty in art as always tainted by human vanity – the vanity of owning artworks, or creating them, or just displaying one's social skills by conforming one's estimations of them to prevailing fashions and the opinion of others – and once again, many of the idiosyncrasies in artistic taste on which people pride themselves fall into this discreditable category. That is why it is only the appreciation of *natural* beauty, *not* beauty in art, that Kant considers the "mark of a good soul." Despite all this, however, toward the end of his discussion of aesthetic judgment in the third *Critique*, Kant does offer a discussion of beauty in art and several topics closely related to it, a discussion that was to prove very influential in the history of modern aesthetics.

'Art' (*Kunst*) in general refers to the capacity of human beings to make things (KU 5:303–304). Many arts are directed at useful objects – tools, houses, shoes, and so on. Others aim at merely the agreeable, such as the art of telling jokes or making conversation or even arranging table music to enliven the mood at dinner parties (KU 5:305). The only kind of art that is an object of pure aesthetic judgment is *schöne Kunst* – 'fine art' or *beautiful* art (the German term for both is the same).

Fine art is "a kind of representation that is purposive in itself and, though without an end, nevertheless promotes the cultivation of the mental powers for sociable communication" (KU 5:306). Objects of fine art may of course also be useful, as a beautiful building may be to those who live in it, or a beautiful speech to someone who wants to persuade by means of it, or beautiful paintings of flowers or birds, when they convey botanical or ornithological information. But the ends they serve are all represented in a concept, while what makes them objects of fine (or beautiful) art is the way they, like beautiful objects in nature, produce a disinterested and universally valid pleasure, free from any concept, through the mere judging of them, and exhibit a form of purposiveness without an end. Kant's theory of the beautiful in art is developed chiefly through two crucial (and closely related) conceptions, that of artistic *genius*, and that of the *aesthetic idea*.

Human art in general produces objects according to rules. In useful arts, the rules are rules of skill, containing knowledge about how to produce an end, and they must be guided by a concept of the end. Fine art, however, insofar as its product is the object of a pure judgment of taste,

cannot be guided by any concept of the end. It must produce a work of art through a human capacity that defies conceptual articulation and is therefore more like natural purposiveness than intentional human purposiveness. This capacity Kant calls *genius*, "the inborn predisposition of the mind (*ingenium*) *through which* nature gives the rule to art" (KU 5:307). It is "a talent for producing that for which no determinate rule can be given" (KU 5:307). Its products are "exemplary" (or "classical") – they serve to others as standards for judging, and also for imitation. But genius cannot be taught, nor its operation explained by those who possess it. It is a kind of inborn gift to those fortunate enough to have it, which defies description or explanation.

In the accounts of genius that stand under the influence of the Kantian one, emphasis is most often placed on the special excellence possessed by the person of genius, as though such a person's gift is to be accounted somehow superior to all other talents of the mind that might be acquired by being taught them, and the grounds and rules for which might be formulated in concepts and words. It is therefore noteworthy that although Kant regards artistic genius as a valuable and unique kind of capacity, he does not regard it as superior to other talents of the mind. He argues, for example, that it is out of place to describe Newton, for example, as a 'genius', since Newton can give reasons for all the scientific propositions and conclusions that form his natural philosophy. "Newton could make all the steps that he had to take, from the first elements of geometry to his great and profound discoveries, entirely intuitive not only to himself but to everyone else, and thus set them out for posterity quite determinately" (KU 5:309). Kant thinks this feature of the great scientist, mathematician, or philosopher makes their ability stand not lower but higher on the scale of what we should value. For what is capable of being communicated, shared, and held in common by rational beings necessarily possesses greater worth than what must separate human beings from one another. The value of all rational beings is absolute, therefore equal. Those human talents are to be most prized which draw people closer to one another and enable them to communicate and share their experiences, cognitions, and practical ends. Genius does this, to be sure, at the level of aesthetic feeling. But it is a greater thing still to be able to do it on the level of common conceptions, as in the science whose rational grounds are open to all, or in moral principles that are universally valid and aim at the universal community of all human beings as a realm of ends.

Thus in the capacity of the scientist to educate and perfect the cognition of others "lies the great advantage of such people over those who have the honor of being called geniuses: since for the latter art somewhere comes to a halt, because a limit is set for it beyond which it cannot go, which presumably has also long since been reached and cannot be extended any more" (KU 5:309). This also exhibits Kant's adherence

to eighteenth-century classicism in aesthetics. In his view, art represents a field of human endeavor whose essential possibilities have already been exhausted. Thus its finest products stand for us as examples to be admired, ever newly appropriated, and even imitated, but never to be fundamentally surpassed. This is connected to the fact that fine art stands essentially lower on the scale of human endeavor than science, morality, or philosophy, which are guided not by natural genius but by reason, whose strivings and capacities are in principle inexhaustible.

Kant's entire world-view is thus at the opposite pole from that of the romantic aesthete who celebrates art because the genius is a "special person" with a quasi-divine gift, whose wondrous products are to be objects of uncomprehending admiration to common mortals. To view the artist in this way is to turn genius into a human gift of fortune like wealth, power, or honor – that is, merely another pretext for human self-conceit. It is a view of art that would transform enjoyment of the beautiful into another pretext for human vice, and helps to show why Kant thinks it is only the appreciation of natural beauty, not of artistic beauty, that is a reliable mark of a good soul.

Aesthetic ideas

The genius that makes fine art possible, however, does make contributions to our cognition that are unique and unlike anything that is possible for our other capacities. Chief of these is the essential product of that faculty of the mind that Kant thinks most constitutes genius, namely "spirit" (*Geist*). Spirit, however, is "nothing other than the faculty for the presentation of *aesthetic ideas*" (KU 5:314).

An idea of reason is a concept, such as that of God, or a simple, indivisible substance, or an uncaused cause, to which no sensible intuition can ever be adequate. An aesthetic idea is just the reverse: it is a sensible representation to which no concept can ever be adequate. Just as in an idea of reason the pure concept *transcends* or *exceeds* our sensory capacities to represent anything corresponding to it, so in an aesthetic idea our sensory imagination represents something that exhausts and goes beyond our power to form any concept capable of comprehending it. "By an aesthetic idea," Kant says, "I mean that representation of the imagination that occasions much thinking though without it being possible for any determinate thought, i.e. *concept*, to be adequate to it, which, consequently, no language fully attains or can make intelligible" (KU 5:314). Aesthetic ideas are images (presented visually, or through other senses, or in words) through which works of art stimulate a thinking process that seems both endlessly fascinating and inexhaustibly suggestive.

Among the examples Kant provides are two attributes associated with classical pagan deities: "Jupiter's eagle, with the lightning in its claws"

and "the peacock of [Juno,] the splendid queen of heaven" (KU 5:315). When we think of the eagle as signifying Jupiter, we are drawn to think of the bird's powerful, curved beak, its terrible scowl, its majestic wings in flight, soaring among the clouds. Juno's peacock makes us think of its proud bearing, the slow, dignified pace of its walk, the shimmering iridescent beauty of its deep, quietly magnificent blue plumage. These thoughts, Kant says, "do not, like *logical attributes*, represent what lies in our concepts of the sublimity and majesty of creation, but something else, which gives the imagination cause to spread itself over a multitude of related representations, which let one think more than one can express in a concept determined by words" (KU 5:315).

It is not accidental that Kant's examples are classical in provenance, but equally non-accidental that they have to do with deities. For just as ideas of reason purport to represent supersensible objects, so Kant seems to think of aesthetic ideas as our most appropriate way of making supersensible objects intuitive or sensibly present to us, particularly objects charged with practical or moral import, such as religious ones. Religion is better represented by the symbolism of beauty than by the terror of the sublime, because its proper function is to uplift us morally toward the ideal, not to frighten us with lurid and superstitious visions of arbitrary divine power administering eternal punishment.

It is not difficult, therefore, to see Kant's theory of aesthetic ideas as closely related to his thesis that beauty is a symbol of morality. The creations of fine art, the products of genius, are sensible representations whose merit for aesthetic judgment is closely allied to their capacity to give a kind of sensuous expression to moral or religious ideas that properly speaking transcend the capacity of our senses to represent them. Further, and from another point of view, aesthetic ideas give a broader significance to some of the elements of Kant's theory of beauty and taste, such as the notion of purposiveness without an end, and the thesis that beauty involves a harmony between imagination and understanding that is free from any concept. The aesthetic idea consists of an infinite wealth of thoughts and associations, unified around a single representation of the imagination, yet with no concept capable of comprehending the unity. This unity is a purposive one, but with no possible concept to represent the end to which they are directed. The aesthetic idea is an image related to sense, but one that suggests an endlessly rich succession of thoughts that the understanding grasps, without a concept, as constituting a harmonious unity.

In both halves of the *Critique of the Power of Judgment*, Kant provides a response to those critics who see philosophical difficulties – or, more often, symptoms of an unhealthy alienation in life-attitude – in his distinctions between understanding and sense, theoretical and practical reason, duty and inclination, the sensible world and the intelligible

world; and also to those (who are often the very same critics) who charge that Kant's philosophy is excessively rationalistic, lacking a proper appreciation for the importance of feeling in human nature and human life. The response is not to deny the reality of the distinctions, but rather to show how human nature allows for them to be mediated, and that precisely through the intervention of feeling. For it is above all in aesthetic experience that we experience the harmony of understanding and sensory imagination, and in the experience of the beautiful and sublime that morality and the supersensible become matters of human feeling for us.

The most general aim of the third *Critique*, to bridge the gulf between theoretical understanding and practical reason, is achieved in aesthetic judgment through coming to see beauty as a symbol of morality and sublimity as an experience of the loftiness of our practical vocation as free beings. It is also achieved in the conclusion of the work, the methodology of teleological judgment, by showing how sensible nature in the organic realm can be regarded as a system of natural ends, and then how that system can be brought to completion only through regarding human beings, who are for morality ends in themselves, as the ultimate end that unifies the teleological system of nature precisely by setting a final end – an end to which all others are ordered and subordinated – in accordance with the laws of morality. Perhaps this is why it was the third *Critique* more than any of Kant's other works that provided inspiration for his idealist followers, Fichte, Schelling, and Hegel.

Yet the *Critique of the Power of Judgment* is above all true to Kant's *critical* enterprise in philosophy. For as it overcomes dualisms, builds bridges, and mediates oppositions, it also carefully preserves the limits of human faculties, by insisting that natural teleology is only a regulative principle of judgment, not a dogmatic doctrine of natural science, and resisting the temptation to see in aesthetic inspiration or artistic genius some occult mode of cognition giving us access to supersensible reality. Perhaps that is why Kant's idealist successors, despite their enthusiasm for this work, never truly understood it and could not accept its solutions to the problems they thought were insoluble for Kantian philosophy.

further reading

Henry Allison, *Kant's Theory of Taste*. New York: Cambridge University Press, 2001.

Ted Cohen and Paul Guyer (eds.), *Essays in Kant's Aesthetics*. Chicago: University of Chicago Press, 1982.

Paul Guyer, *Kant and the Claims of Taste*. 2nd edition. New York: Cambridge University Press, 1997.

— *Kant and the Experience of Freedom*. New York: Cambridge University Press, 1993.
— (ed.), *Kant's Critique of the Power of Judgment: Critical Essays*. Lanham, MD: Rowman and Littlefield, 2003.
Salim Kemal, *Kant and Fine Art*. Oxford: Clarendon Press, 1996.

notes

1 Paul Guyer, *Kant and the Experience of Freedom* (New York: Cambridge University Press, 1993).
2 Kant defines 'imagination' as the faculty through which we represent an object that is not itself present in intuition (KrV B251). We reproduce in imagination the Eiffel Tower which we have seen, even though we are not seeing it now. By combining such remembered images as horses and wings, we imagine a horse with wings even though no such horses exist to be intuited. Our 'productive' imagination is what enables us to intuit objects that require a synthesis through time to be fully present to us. By extension, it is imagination that enables us to experience something in a mode where we are indifferent to whether it actually exists or is merely being represented.

chapter 9

politics and religion

In the Preface to the *Critique of the Power of Judgment* (1790), Kant wrote that with this book "I bring my entire critical enterprise to an end" (KU 5:170). Kant may have ended the enterprise of critique in 1790, but he continued to write and publish for the next eight years, developing and completing parts of the system of philosophy for which he had tried to lay the foundations in his critical works. Kant's most significant publications during the last decade of his life dealt with practical questions of universal human concern – with politics and religion.

I the concept of right[1]

In Anglophone philosophy, "Kantian" views on political philosophy seldom orient themselves to Kant's own writings on that subject. Instead, they project what they think is implied about politics from Kant's moral philosophy – chiefly as it was expressed in his most abstract and foundational treatise, the *Groundwork*. Kant's authentic political philosophy therefore remains much less well known than one might have thought. Part of the explanation for this is the fact that Kant's political thought was expressed only late in his career, and all too obscurely at that. Even fundamental issues, such as Kant's conception of the relation of political to moral philosophy, remain unclear and matters of controversy among those who have studied them.

One important question is whether, or in what sense, the two parts of Kant's *Metaphysics of Morals*, the Doctrine of Right and the Doctrine of Virtue, are really parts of a single doctrine, falling under a single principle. This is a difficult question, and in chapter 7, I have already argued for a negative answer to it. The theory of ethics is a theory about any human being's regulation of its own conduct in accordance with self-given laws of reason. The theory of right is a theory about the rational standards for externally coercive laws and the foundations of the human institution (called 'civil society' or 'the political state') within which such laws have their place. It is an important tenet of Kantian doctrine that ethical

duties are laid on each person autonomously by that person's own reason, that the proper incentive for their fulfillment is the person's own inner motive of duty, and that it is wrong and improper for others, or for society in general, to attempt to compel us to fulfill them. Duties of right, by contrast, are essentially imposed from outside the agent by an external power, and the justice or rightfulness of the actions that fulfill them is the same whatever the motive – that is, payment of a debt or obedience to laws against theft are equally just whether they are motivated by a sense of duty or by immediate fear of what a court or a policeman will do to you. The sphere of right derives the concept of duty from the moral imperative (MS 6:239), but it does not follow that this is also the ground of imperatives of right. Right and its externally enforced laws constitute a closed system within themselves, even if Kant also thinks the system as a whole can be given rational support from outside it by moral principles and rational beings *also* have an ethical duty to fulfill duties of right.

The point of the separateness of the spheres of right and ethics can be brought out by issuing a Kantian challenge against the legitimacy of two related distinctions that tend to be taken for granted by moral and political philosophers, especially in the Anglophone tradition. The first distinction is between moral and legal rights, the second between morality and positive law. We sometimes say that a person has a right to something, even though we know he lacks a right to it under existing laws. But such talk is ambiguous, and can refer to two very different kinds of cases. If someone has an obligation to do something for me (even a purely moral one, not one that could ever be enforced under the law), then in one sense of 'right' it follows that I have a *right* to that performance (and if this is not a right that is, or even should be, enforceable under the law, we naturally call it a 'moral right'). In other cases, however, we think people *should* have rights enforceable under the law that at present they do not have because existing laws (we think unjustly) do not recognize these rights. These too we call 'moral rights' (to contrast them with the rights people may claim under the laws as they are). If we think these two cases are alike, then we are in danger of thinking that all duties are such that it is appropriate to enforce them through external coercion, or else that the only rational standards by which laws may be judged correct or incorrect are *moral* standards.

Some people may actually believe one or both of these things, but Kant fundamentally disagrees with both. He thinks all properly *ethical* duties – duties of beneficence to others, or duties to perfect oneself by developing one's talents – are such that it would be not only morally impermissible but even contrary to right to attempt to enforce them through coercion (as by passing laws punishing those who fail to fulfill them). He also thinks that since morality's concern is with the promotion

of human perfection and happiness through voluntary conduct motivated by autonomous reason and duty, the proper standards to which to hold the coercive laws of the state are not moral standards, but rather standards proper to right, standards geared solely to safeguarding the external freedom of rational beings.

This is why it is also inappropriate to think that when we criticize existing (or positive) laws, we should do so by appealing to *morality*. From a Kantian point of view, there is even something illegitimate about the notion of 'positive law' as used by theorists who try to reduce claims about 'positive' law merely to factual or causal claims about what will occur under some presently existing system of legal statutes and institutions. Kant shares with the natural law tradition the view that to talk about law (or right) at all is always to speak normatively, not merely to report what will happen but to say what ought to happen according to a set of norms that conform at least minimally to certain rational standards. The proper standards, in Kant's view, however, are not moral or ethical standards, those appropriate to the general regulation of the conduct of a rational being as a moral agent, but rather the standards of right, those appropriate to the regulation of a social system of coercion – a political state with its systems of civil and criminal law.

II the system of right

For Kant, the system of right begins with the one innate right every human being possesses simply in virtue of their humanity or rational nature – the right to freedom, or independence of being constrained by another's arbitrary will (MS 6:237). To this right belongs also the right of *equality* – immunity from being bound by others to more than one can bind them – the right of being *one's own master*, and the right of being "beyond reproach," that is, considered to have done no wrong to others as long as you have not done anything to diminish what is theirs by right.[2] Kant divides our basic duties of right into three headings, based on formulae used by the Roman jurist Ulpian: *honeste vive, neminem laede, suum cuique tribue*. Under the first of these, "live honorably," Kant understands asserting one's own worth as a human being, not turning yourself into a mere means for others. That this juridical duty or duty of right is distinct from our ethical duty of self-respect is indicated by the fact that Kant proposes to derive it from the *right* of humanity in our own person (MS 6:236). The second formula, "injure no one," Kant understands to obligate us not to live with others except under conditions of right. Later Kant will argue that this duty requires us to leave the state of nature and enter into the civil state, and authorizes every person to use coercion to force others to do this as well (MS

6:306–312). The third formula, "give to each what is his," Kant considers to be an empty tautology unless it is understood in this sense: "Enter a condition in which what belongs to each can be secured to him against everyone else" (MS 6:237).

The first main part of the Doctrine of Right, Private Right, is concerned with the conception of "what is mine or yours," that is, with the foundation of rights of property. Kant distinguishes between two kinds of possession, which he calls 'phenomenal' possession and 'noumenal' (or 'intelligible') possession. I possess an external thing phenomenally when I am in immediate bodily contact with it (for instance, holding it in my hand). It is obvious how I am wronged (my external freedom is violated) when something in my phenomenal possession is taken from me against my will, because this involves a physical violation of my body. Kant argues, however, that people cannot carry out their free projects unless they can also be wronged through the removal of or interference with external objects that are not in the immediate physical control of the owner, but in the owner's possession only through a pure concept of the understanding. For this reason, we must postulate that this kind of noumenal or intelligible possession is also possible (MS 6:249–252).

The primary form of property, Kant argues, is property in land, since this is a precondition of appropriating other things that are found or made on the land. Fundamentally, however, Kant thinks, both the land and all things on it are in the common possession of each of us with all others (MS 6:261–262). Thus my intelligible possession of any thing is based on the idea of a legislative act of all giving me rightful possession of it (MS 6:268). Based on such rights of property, Kant derives the notion of rights by contract (MS 271–276) and also rights over the status of persons who are related to you as husband, wife, child, or domestic servant (MS 6:276–286).

An act of all according noumenal possession of an object is ideally thinkable even in a state of nature, that is, apart from, or in abstraction from, a civil society determining and enforcing right. But since in the state of nature there is no common judge of right to settle possible disputes that might arise about who owns a given thing, such possession is always only "provisional," never "peremptory" – that is, enforceable against those who might dispute it. Thus genuine and enforceable rights of possession are possible only in a "rightful condition" or "civil condition," in which there is an "authority giving laws publicly" (MS 6:255–256).

For Kant, as for other modern political theorists such as Locke and Rousseau, the original and fundamental purpose of the political state is to define and enforce rights of private property. But since peremptory rights of property are held only in a rightful civil condition, and this requires subjection to a legislative authority, Kant's theory treats the state as the "supreme proprietor" of all land and other property (MS

6:323–325). This is the basis of Kant's argument that the state has the authority to tax the wealthy for the support of the poor (MS 6:325–326). The wealthy have no right to complain about this, because they not only owe their own existence to the protection of the state, but their right to own whatever they have is also conditional on the laws, including those providing for the taxation of their wealth for the benefit of the poor.

Right in a civil condition or state is the subject of the second main part of the Doctrine of Right, Public Right. Kant accepts the fundamental division of powers within a state derived from Montesquieu, namely, that into legislative, executive, and judicial authorities (MS 6:313–316). He understands the division between legislature and executive, following Rousseau, as that between the power that lays down general laws and a power that forcibly commands them to be obeyed in particular cases; the judiciary's function is to apply the law to the particular case. (Kant thus compares the three authorities to the major premise, minor premise, and conclusion of a syllogism, MS 6:313.) He also insists that these two functions cannot be combined in the same person or group of people. For this reason, he rejects as wrongful a "despotic" government, which gives the laws as well as enforces them.

In Kant's view, the only constitution that truly accords with right is one involving the separation of legislative from executive powers guaranteeing the equal right of all citizens, and in which the legislature and government are representatives of the people (EF 8:352). This constitution is that of "a pure republic," a " representative system of the people, in order to protect its rights in its name, by all the citizens united and acting through their delegates (deputies)" (MS 6:341; cf. EF 8:349). However, Kant considers a "democratic" government, in which the executive power belongs to the mass of the people, the most dangerous and the most conducive to despotism (MS 6:339, EF 8:351–353). A state that is not republican in its constitution (such as the Prussian state Kant lived in all his life, which was an absolute despotism in which all important political and military offices were restricted to a hereditary nobility) can rule legitimately, but only if it legislates in a republican spirit and rules in a manner that leads it in the direction of reforms leading to a republican constitution (EF 8:352–353, cf. MS 6:340, 370).

Kant accepted existing occupational and property qualifications for voting and officeholding that held in his day wherever such institutions existed at all. Only those who are not economically dependent on others, he argued, are in a position to give their independent voice, and participate in the state as "active citizens", by voting and holding office. The rest (including servants, wage-laborers, peasants working land owned by others, and of course all women) are "passive citizens": the state protects their rights, but they have no claim to participate in making decisions for it.

One of Kant's most famous (or infamous) doctrines about the state is that it is always wrong to disobey even unjust laws, or the commands of an unjust ruler acting contrary to law (as long as these laws and commands do not require you to do something that is in itself morally wrong), and always wrong to overthrow the existing ruler, however unjustly that ruler may behave (MS 6:371–372, TP 8:297–305). Kant reconciled these doctrines with his admitted enthusiasm for the French Revolution and his equally enthusiastic support of the French Republic through the (extremely dubious) argument that Louis XVI had not been overthrown by force but had *voluntarily* abdicated "to extricate himself from the embarrassment of large state debts" (MS 6:341). The experience of Kant's century, however, was that progressive reforms occurred for the most part only when monarchs (such as Frederick the Great, Catherine the Great, or Joseph II) were persuaded to make them. His principled arguments against rebellion and insurrection even against an unjust ruler convince few people today. They depend on the rather Hobbesian idea that a condition of right itself depends on there being a head of state to whose commands the unconditional obedience of all is the fundamental demand of right. Kant did not carry this idea to the Hobbesian extreme of saying that citizens have no rights at all against the head of state. He argued rather that although such rights are in principle unenforceable against the ruler, they are real, and a ruler who respects the idea of right will not violate them (TP 8:289–297).

More generally, Kant's theory places principled restrictions on both rulers and acts of legislation if they are to be considered just and to accord with the idea of right. It is in this aspect of Kant's doctrine above all that we find the norms governing people's coercive action on one another that constitute the standards of right that are distinct from (though of course in harmony with) the standards of morality. These restrictions include there being a civil constitution that guarantees the *freedom*, *equality*, and *independence* of citizens (MS 6:314, TP 8:290). It also includes an ideal restriction on legislation, "the idea of an original contract." Kant does not regard the state as literally based on a contract, but he thinks that the idea of the people as giving unanimous rational consent to a system of laws can function as a way of distinguishing just laws from unjust ones. A legislator, that is, should give only such laws as could have arisen from the united will of the whole people that are subject to them (TP 8:297, MS 6:340). Kant thinks this would forbid, for example, a church from binding itself forever to articles of faith or to religious practices which would make all future progress or enlightenment in religion impossible (TP 8:305). He also argues that it forbids the institution of restricting all high political offices to a hereditary nobility (TP 8:292–294, cf. EF 8:350–351, MS 6:329), as was always the case in the state in which he lived. Kant also argues that standards of right subject

the acts of heads of state to two "principles of publicity," a negative one and an affirmative one: "All actions relating to the right of others are wrong if their maxim is incompatible with publicity" (EF 8:381) and "All maxims that *need* publicity (in order not to fail in their end) harmonize with right and politics combined" (EF 8:386). Both principles can be seen as furthering the idea of the original contract, since they require rulers to conduct themselves openly before citizens in ways that would be required for there to be unanimous consent to what they do.

III the right of nations and perpetual peace

Some of Kant's most original and forward-looking contributions to the theory of right lie in the area of international relations. In 1713, at the time of the Peace of Utrecht, the Abbé de Saint-Pierre put forward what was then universally regarded as a utopian project for what he called a 'European Union', an organization of states for the maintenance of peace in Europe. In 1795, on the occasion of the Treaty of Basel between France and Prussia, Kant wrote his treatise *Toward Perpetual Peace*, proposing a federation of nations, perhaps beginning in Europe but in principle spreading to all nations of the earth, whose aim was to eliminate not only war but also the more or less permanent preparations for war, which Kant thought of as warping the collective efforts of humanity toward a future worthy of human dignity.

Perpetual Peace was the definitive expression of ideas Kant had articulated in *Idea for a Universal History* over a decade earlier, and had also expressed in the third part of the essay on theory and practice in 1793. It begins with six "preliminary articles" concerning the conduct of states in war or in preparations for war, that are designed to make a permanent condition of peace between them more likely. These include refraining from interference in the constitution or government of another state, the gradual elimination of standing armies, and the contracting of no national debts for the purpose of making war. Then Kant states three "definitive" articles providing for a condition of perpetual peace between nations. The first states that the constitution of every state shall be that of a republic, since wars will be less likely if the governments who must make them represent the people who must fight them and bear the cost of them (EF 8:349). The second article proposes a federation of free states maintaining peace and justice between all of them (EF 8:354). The third article deals with the conditions of "universal hospitality" regarding the treatment of citizens of one state when they visit another (EF 8:357). The three articles are constructed corresponding to three kinds of right that were traditionally distinguished: the right of citizens in a state (*ius civitatis*), the right of nations in their relation to one

another (*ius gentium*), and the right of citizens of the world as human beings (*ius cosmopoliticum*) (EF 8:349).

Under this last heading, Kant discusses the issues of cosmopolitical right that are raised by the practices of European colonialism – toward which his attitude is one of unqualified disapproval. Though he thinks European civilization is more advanced than that of other parts of the world, Kant regards civilization itself as directly unfitting Europeans for the gratuitously self-appointed task of civilizing others.[3] Civilization, in his view, is a process through which any nation or people must raise itself to greater powers by its own actions, guided by its own judgments and aims. Kant regards greater civilization as developing the capacities of our species, and thus as having the potentiality for eventually leading people toward a better life in matters of both right and morality. But in the comparatively early stages of it that we find among people now, the chief marks of civilization are the sorts of tyranny, greed, ambition, and duplicity that bring corruption and inequality to European society itself, and lead Europeans arrogantly to seize territories in other parts of the world, looking upon them "as belonging to nobody; for they counted the inhabitants as nothing." Kant speaks of the "horrifying lengths" to which Europeans have gone in "visiting" foreign lands – "which with them is tantamount to conquering them" (EF 8:358). It is easy, he says, to see through the "veil of injustice" by which colonizers sanction their violations of right by appealing to the good ends they propose to achieve by it (MS 266). One might at least have hoped that the lesson Kant is teaching here could have been learned sometime in the two hundred years since he wrote. But recent events (I mean the 2003 American and British invasion of Iraq) prove that our backward leaders still flatter themselves that they are doing noble deeds when they inflict their wrongful imperialist wars on non-Western nations. (Only the hypocritical verbiage has changed – from 'civilizing' those they conquer, they now aim at 'liberating' them. In fact, however, they are equally incapable of doing either – which is hardly surprising in the case of the American instigators of this imperialism, since they represent everything that is least civilized and most hostile to liberty in the political culture of their own country.)

No nation, in Kant's view, has the right to settle, still less to invade militarily, the territory of another without a specific contract with the indigenous people permitting them to be there. This is not a matter of philanthropy or morality, but a fundamental matter of *right* (MS 6:353). Equally impermissible by right are the fraudulent contracts through which Europeans have imposed their rule over peoples in America, Africa, and Indonesia – "becoming owners of their land, making use of our superior force without regard for their first possession" (MS 6:266). Kant therefore endorses the wise policy of the Chinese and Japanese in

limiting the access of Europeans to their territories, or at least "excluding them, like prisoners, from community with the natives" (EF 8:359).

Following the 'articles', Kant provides two 'supplements'. In the first he argues for the possibility (though by no means the inevitability) of perpetual peace based on the philosophy of history. In the second supplement and in an appendix added to the second edition of the essay a year later, Kant makes a philosopher's plea to politicians. He argues that in conducting affairs of state they should follow the maxims of philosophers, and subject their politics to principles of right. This is the only way, Kant argues, that even their prudential interests are likely to be served in the long run.

Kant here willingly casts himself in the role of the unpragmatic moralist or philosopher, offering advice or even admonition to the cynical politician. This doubtless reflects in part his uneasy relationship with the Prussian authorities since the death of Frederick the Great, and especially since Wöllner's letter of reprimand to Kant in 1794. But down to this day it is common for some politicians to see themselves as "realists," and look with contempt on the "ivory tower philosophers" who foolishly hope to "make the world a better place." But since it is a characteristic of human nature that the same traits that make people greedy also make them shortsighted, it is also still predictable that the unprincipled ambition that goes under the names of "realism" or "pragmatism" frequently leads politicians into disasters which would have been avoided if only they had listened to the impractical philosophers.

IV moral faith and religion

Kant's basic attitude toward religion was typical of an Enlightenment thinker, especially a thinker of the German Enlightenment. He was deeply suspicious toward popular religious beliefs and practices, and hostile toward clerical power, both in politics and over people's minds. But he was not hostile toward what he understood to be *true religion*; on the contrary, he regarded it as extremely important. His attitude toward religion could be compared in this respect to his attitude toward the state. Political institutions – as they have been, and as they are – represent mainly tyranny and injustice; they are to blame for terrible social inequalities and for the wars, and preparation for war, that stifle human potentiality and stand in the way of progress. But the proper function of the state – the coercive protection of the right of rational beings to external freedom – is indispensable to human life. Without it neither culture nor moral progress would be possible.

Analogously, past forms of religion have arisen out of superstitious fear, a slavish cast of mind, and the ruthless ambitions of priests to

subject even the inner life of human beings to their tyrannical tutelage. But the proper function of religion is to bring human beings together for the purpose of the collective moral improvement of the human race. In Kant's view, we can no more expect to fulfill our vocation as moral beings apart from religion than we can expect to achieve justice through anarchy. The essential thing, then, in both politics and religion, is to reform, through free communication and enlightenment, the everyday thinking of people about matters of right and religion – so that these corrupt and oppressive institutions can be made capable of doing what reason calls upon them to do in furthering external freedom and moral progress.

The moral argument for God's existence

Kant famously declared that his reason must deny knowledge in order to make room for faith (KrV B xxx). He held that although theoretical reason can provide us with no cognition of God, no proofs of God's existence, practical considerations can justify a belief, at least for the purposes of moral action, that there is a wise, benevolent, and just Providence ordering the world.

Since reason always seeks the unconditioned, Kant argues, as a moral agent under the guidance of reason, I will form the conception of a single goal unifying the object of my efforts and those of other well-disposed moral beings. The traditional name for this goal is the "highest good" (*summum bonum*). Kant maintains that the highest good has two distinct components, morality and happiness. Morality, goodness of will and of moral conduct, has to do with the worth of our person, happiness with the worth of our state or condition. Morality is unconditionally good, but happiness is good only under the condition that those who enjoy it have made themselves worthy of it through the goodness of their will and conduct. Regarding myself, the highest good consists therefore in achieving the morally best character and volition I can, and then enjoying happiness in proportion to my worthiness to be happy. The highest good for the whole world would be a world of virtuous agents who enjoy happiness in proportion to their worthiness.

Now Kant also holds that it is rational to pursue an end only insofar as you believe the end possible of attainment through the actions you take toward it. It follows that if you believe that some proposed end is impossible of attainment through any actions open to you, it is contrary to reason to set that as an end. The question then naturally arises whether the highest good is possible of attainment through the actions that I and other well-disposed moral agents might take toward it. Regarding the happiness component of the highest good, it does not seem that the laws of mechanical causality governing nature are such as to insure that the

happiness of moral beings will be proportional to their worthiness. Nor does anything else we know about the natural world through experience give us any grounds for believing that the highest good is possible of attainment through our moral strivings. We cannot show that the highest good is impossible, but we also have insufficient reason for thinking it is possible. Yet as rational moral beings, we ought to make the highest good our end. This leaves us in a kind of perplexity regarding our relation to the world in which we must act. Morality requires us to set an end about which theoretical reason gives us insufficient grounds for believing that it is possible of attainment. We are threatened with an incoherence between our practical volition and our justified beliefs and assertions about the world.

If we resolve the problem by changing our volitions, and ceasing to make the highest good our end, then we are abandoning our commitment to something morality tells us to do. Thus we should not resolve the problem in this way. But the only other way to resolve it is to change what we affirm or believe about the world. This gives us a reason, deriving not from theoretical evidence but solely from practical considerations, for affirming – at least relative to our practical strivings for the highest good – that the world is so ordered that happiness is somehow proportioned to worthiness and so the highest good is possible of attainment through our efforts to make better people of ourselves and to achieve the various finite ends that morality proposes for us. We can best give content to this belief in the attainability of the highest good by supposing that there is an omnipotent, omniscient, and morally perfect being who orders the world according to its benevolent and just will. The idea of such a being – an *ens realissimum* or God – has already been shown to be an indispensable ideal of theoretical reason, even if theoretical reason is forever unable to demonstrate either the existence or the non-existence of its object. The only reasonable way to resolve our practical problem about the possibility of the highest good is to go beyond what theoretical reason can affirm about this idea, and assent to the existence of its object.

There are some variations among Kant's writings in the way this line of argument is presented. In the *Critique of Pure Reason*, our hope for the happiness of which we have made ourselves worthy is presented as a necessary *motive* for acting morally, but after the *Groundwork*, only duty (respect for law or for rational nature as an end) is seen as a worthy motive for moral conduct. In the first *Critique*, the happiness we hope for is also located in another life, but this idea disappears in the second and third *Critiques*, and in the *Critique of the Power of Judgment*, happiness in proportion to worthiness (now called 'the human being under moral laws') is seen as a final moral end that is to provide the teleology of nature with its unification (which would seem to imply that we are to

believe that the highest good is realized entirely within nature, not in a supernatural afterlife).

There are also some subtle variations in Kant's statements of the conclusion of the moral argument, which relate to some serious questions about how strong a conclusion the argument establishes. If the basis of the argument is that we must hold that the highest good is *possible*, and the existence of an all-powerful, benevolent, and just God would guarantee that the highest good actually comes about, then it would seem that making the highest good an end does not rationally commit us to the actual existence of God, but only to the *possibility* that there is a God. Kant sometimes seems to agree to this, insisting that the "minimum of theology" needed for religion and moral conduct is not an "assertoric faith" but merely the belief that God *possibly* exists (R 6:154, cf. Ak 28:998). But it is equally evident that Kant thinks assent to the actual existence of God is better suited to a properly moral disposition than is this "minimum."

There is also a question whether the argument really justifies a full-blown *belief* (or faith, *Glaube*) in God (as Kant sometimes claims for his argument). Reflecting on the fact that, for practical purposes, assenting to the existence of God removes a rational conflict between our belief and purposive action cannot by itself produce belief in God, any more than being offered a large sum of money if you believe that human beings are descended from space aliens can directly produce the profitable belief. Actual belief can come about only through evidence or else through contra-rational processes such as the Jamesian "will to believe" – in other words, wishful thinking, self-deception, or biased attention to weaker evidence over stronger evidence (shameful and degrading self-manipulative devices, not uncommon in popular religious thought and practice, but never mentioned by Kant and surely not countenanced by him). Reflection on the paradox of what our pursuit of the highest good presupposes might, however, lead to a rational *assent to* or *acceptance of* the proposition "God exists" for the *practical* purpose of resolving the paradox. It often makes sense, in science for instance, to speak of accepting a hypothesis in certain respects or for certain purposes (for example, as part of a strategy of inquiry or for heuristic purposes). But it would seem odd, or out of place, to speak of *believing* things only "in certain respects" or "for certain purposes."

Kant does often speak explicitly not of belief (*Glaube*) but rather of "assent" (*Fürwahrhalten*) or "acceptance" (*Annehmen*) of the existence of God "in a practical respect" or "for practical aims" (*in praktischer Absicht*). No doubt Kant wanted to think of his moral argument as providing a rational basis for the heartfelt (and morally engaged) attitude that religious people call their "faith" in God. But it is questionable whether his moral arguments can really deliver what religious people

want here. This may be simply because no philosopher who denies us proofs and evidence of God's existence, as Kant does, can honestly condone an actual *belief* in God.

Religion and the ethical community

Human beings, according to Kant's theory of history, have a natural tendency to "self-conceit" or "unsociable sociability." They seek superiority over other human beings, and are drawn into society with others more by competitive impulses than by common interests. But moral reason tells them that other human beings are their equals in dignity as ends in themselves, and commands them to live by the laws of a realm of ends, laws that command them to seek unity among human ends rather than self-seeking and competition. The natural tendency to competitiveness, therefore, amounts to a fundamental maxim of placing selfish ends ahead of the commands of morality and constitutes what Kant (in *Religion Within the Boundaries of Mere Reason*) characterizes as the "radical propensity to evil in human nature" (R 6:29). Both because the moral law commands us to bring our ends into unity with those of others and because action in isolation from others is the fundamental way in which our propensity to evil manifests itself, Kant argues that human beings are incapable of making progress in improving their moral characters if they do so individually, each of us privately engaged in an inward struggle with our own inclinations. The task of morally improving ourselves, both as individuals and as a rational species, must lie rather in participation in a certain kind of society that is devoted to combating the radical propensity to evil and furthering the ideal of a realm of ends on earth. For this reason, the future of humanity depends on the success of a certain kind of communal enterprise, to which Kant gives the name "ethical community" or "ethical commonwealth" (*ethische gemeine Wesen*) (R 6:97).

Just as the historical model for institutions protecting external right is a political state and its system of coercive laws, so the historical model for the ethical community is organized religion – a church or ecclesiastical community. Yet just as the existing political state stands in need of fundamental reform if it is to fulfill its rational function in human life, so too religion as it has been, and is, remains far from what it needs to be if it is to contribute as it should to the moral improvement of the human species. Religious communities have usually been founded on a supposed divine revelation, typically in the form of some scriptural document which has been accepted as authoritative. They have usually been ruled by a class of priestly tyrants, who have done more to enslave than liberate the mind and spirit. The idea of serving God that such communities have had has often been corrupt and superstitious, consisting of a

set of morally indifferent or even degrading constraints on conduct (the performance of rituals, meaningless restrictions on what people eat or when they are permitted to work, regular performances of fetishistic conjurations of divine presence or formalized practices of slavish praise and contemptible begging directed at the divine being – conceived, accordingly, as a vain tyrant who is disposed to favor unjustly those cringing subjects who most flatter him and abase themselves before him). Religious communities have also frequently served more to promote conflict than unity among human beings, as warring peoples think of themselves as serving different and mutually hostile deities. Or else ecclesiastical faiths supposedly worshipping the same single true God murder and enslave each other as each ecclesiastical community, thinking arrogantly that it has some sort of exclusive access to the divine will, tries to impose its own superstitious beliefs and morally superfluous practices on all the others.

True religion, therefore, in Kant's view, would be as different from traditional religion as a just political commonwealth would be from the military despotism of absolute monarchical states (such as the Prussian state under which he lived). Because the ethical community is concerned with the virtuous use of inner freedom rather than with the right to external freedom protected coercively by the political state, membership in the ethical community must be wholly voluntary, and obedience to its laws is not something to which we should be motivated by coercive sanctions of any kind.

Religion, Kant says, is "the recognition of all duties as divine commands" (R 6:153–154). True religion consists in regarding all human duties (given to us by our own autonomous reason) as also legislated by the rational will of a supremely real being or God. Moral faith consists in assenting to the idea that God regulates the world wisely and beneficently under the same moral laws. In this way, the will of God can serve the ethical community as its public legislator, providing human individuals with a common set of (non-coercive moral) laws. But Kant holds that even belief in God's existence is not necessary for religion, since with even the "minimum of theology" (that God is possible) one can think that if there is a God, then one's duties are commanded by God. The true ethical community would not be held together by creeds and catechisms, nor would it involve the "humiliating" distinction between clergy and laity. It would place service to God in doing one's human duties rather than in morally indifferent statutory observances or superstitious attempts to conjure up divine favor for oneself or one's projects.

Kant does not expect existing religious communities immediately to achieve the form of true or rational religion any more than he expects existing political states immediately to assume properly just republican

constitutions. Kant acknowledges that by a weakness of human nature, ethical communities seem to need some scriptural document containing purported divine revelation in order to establish themselves in the world. Kant's "religion within the boundaries of mere reason" does not reject such documents (since reason can no more declare empirical revelation by God to be impossible than it can attest to the validity of any particular claims to divine revelation). Instead, rational religion will undertake the task of interpreting them in such a way as to make them consonant with the idea that they might have been revealed by a supremely wise and morally perfect being – that is, interpreting them so that what they are taken to mean is consonant with morality as our reason best understands it. This means, Kant says, that God cannot be understood to have commanded, for instance, that a man should murder his innocent son just to demonstrate his propensity to blind obedience (R 6:187).

Kant's chief aim in the *Religion*, however, should be seen as a positive one regarding ecclesiastical faith, and especially Christianity. In Kant's view, the moral progress of the human race is possible only through the advancement of religion in fulfilling its proper rational vocation. Accordingly, Kant wants to show how the central experiences of the moral life, involving our struggle against the evil in our nature, our doubts and hopes regarding ultimate victory in this struggle, and our striving for moral improvement both within ourselves and along with others, can find expression in the concepts, doctrines, thoughts, and feelings with which Christians are already familiar from the practice of their faith. It is therefore simultaneously an attempt to provide a progressive and rationalistic interpretation of that faith, and to provide a rational defense of living Christianity by exhibiting the Christian life as an entirely suitable way of fulfilling our moral vocation. Kant even privileges Christianity among ecclesiastical faiths by claiming that it alone "issued from the mouth of its first teacher not as a statutory but as a moral religion" (R 6:167). At the same time, he is quick to note the many ways in which historical Christianity has deviated from the spirit of this original teaching, and he would never deny that other faiths that began as statutory religions are equally capable of coming to be moral religions.

Almost immediately Kant's conception of a moral religion was rejected by the Romantics, perhaps most famously and articulately in Schleiermacher's *Speeches on Religion* (1799), which celebrated what is culturally specific and 'positive' in religion as the only true expression of its essence, regarding the idea of a universal rational religion as a bloodless fantasy of abstract philosophers who had lost their sense for the truly religious. Schleiermacher also scorned Kant's attempt, as he saw it, to reduce religion (the highest thing in human life) to a mere means for

the promotion of morality. If this is the verdict of the father of religious modernism, then we should not expect a more favorable attitude from theological conservatives toward an austere religion of reason that spurns miracles and the *Schwärmerei* of supernatural experiences, and sternly disapproves of all forms of self-abasing worship and of any attempt to replace the always troubled reliance on natural reason with the spiritual security afforded by deference to revealed authorities. Secular moralists, however, are apt to be equally distressed by Kant's attempts to claim that Christian categories are the right ones in which to think about and experience the truths of morality. Such attempts often have only the effect of leading them to question the rational respectability of Kantian ethics itself.

Yet Kant's philosophy of religion was grounded on the historical hope that there would be a convergence between religion and enlightened reason. All our reservations about it must be attributed, in the end, to the sad fact that what he hoped for has simply failed to come to pass. But Kant's hopes for religion, however much they have been disappointed, must be seen instead as the form taken for him by a hope that many of us still share – the hope for the gradual progress of the human species in history toward a realm of ends in which the divisions between people will be overcome and humanity will be united in a cosmopolitan moral community which respects the rights of everyone and unites the happiness of each with the happiness of all as a shared end of human striving. This was the hope with which Kant ended his lectures on anthropology, the last major work published under his own name, and in this sense it may be called literally Kant's last word about the human condition:

> In working against the [evil] propensity [in human nature] . . . our will is in general good, but the accomplishment of what we will is made more difficult by the fact that the attainment of the end can be expected not through the free agreement of *individuals*, but only through the progressive organization of citizens of the earth into and toward the species as a system that is cosmopolitically combined. (VA 7:333)

further reading

John M. Hare, *The Moral Gap*. Oxford: Clarendon Press, 1996.

Alexander Kaufman, *Welfare in the Kantian State*. Oxford: Oxford University Press, 1999.

Leslie Mulholland, *Kant's System of Rights*. New York: Columbia University Press, 1990.

Philip L. Quinn (ed.), *Kant's Philosophy of Religion. Faith and Philosophy*, Volume 17, No. 4 (October, 2000).

John Rawls, *Lectures on the History of Moral Philosophy*, ed. Barbara Herman. Cambridge, MA: Harvard University Press, 2000.

Allen D. Rosen, *Kant's Theory of Justice*. Ithaca: Cornell University Press, 1993.
Philip J. Rossi, S.J., and Michael Wreen (eds.), *Kant's Philosophy of Religion Reconsidered*. Bloomington: Indiana University Press, 1991.
Howard Williams, *Kant's Political Philosophy*. New York: St. Martin's Press, 1983.
Allen W. Wood, *Kant's Moral Religion*. Ithaca: Cornell University Press, 1970.

notes

1 The German term *Recht* means 'right' but also refers to the entire system of law in a state and at the same time to its rational foundation. The study of law in a university takes place in a *Rechtsschule*. This is contrasted with the term *Gesetz*, which means 'law' in the sense of a particular legal statute. The distinction is drawn from the Latin distinction between *ius* and *lex*, and has its equivalent in virtually all European languages except English. (The equivalent of *ius* or *Recht* is *droit* in French, *diritto* in Italian, *derecho* in Spanish, *prawo* in Polish, *jog* in Hungarian, and so forth.) But the closest term in English, 'right', has a meaning that is in one sense much narrower, since it would not be applied to a school of law, for example, and in another much broader, since it can refer to any action that meets any sort of standard of correctness, including, but by no means limited to, moral standards. This linguistic oddity may be one reason why Anglophone philosophy tends to commit the errors I am trying to criticize here.

2 It is noteworthy that under this last right Kant includes the right of communicating thoughts to them, whether what you say is true and sincere or false and insincere, as long as it remains up to them whether they want to believe it (MS 6:238). This might seem to contradict what Kant says in his brief but famous (or infamous) essay "On a Presumed Right to Lie" (Ak 8:425–430), where Kant denies that we have the right to lie even to a would-be murderer who asks us about the whereabouts of his intended victim. There is no inconsistency once we realize that in the discussion of this example, the would-be murderer is being assumed to have a right to rely on what we say, in the way that you might have the right to rely on what I say about the condition of my house when I offer it to you for sale. Once we realize that Kant's discussion of the lie is predicated on this assumption, it is easier to see why he draws the conclusion about it that he does (the conclusion that virtually everyone finds shocking and morally perverse). The puzzling thing then is to understand why Kant is making this assumption about the example, instead of assuming that what I say to the would-be murderer is something he is at liberty to believe or not, as he likes (in which case Kant thinks I have a perfect right to lie to him). There is no space here to try to resolve this puzzle. But readers of the essay on the presumed right to lie will not understand it unless they appreciate the background, a controversy between Kant and Benjamin Constant, which might help to explain why Kant understands this example in such a counterintuitive way.

3 In his lectures on anthropology, Kant even exhibits this assessment of European superiority as grounded on a theory of racial superiority. He regards the human species as biologically one, but thinks that different racial characteristics can be developed by living in different climates and adopting different

modes of life, and then that these characteristics can be passed on to descendants. Of the four races Kant recognizes – (1) White, (2) Asian or "Yellow Indian," (3) Negro, and (4) "copper-red American" – he thinks this list, in this order, represents their respective potentialities for contributing to human civilization (Ak 25:840, 843,1187). It is noteworthy, however, that despite such deplorable views, Kant never suggests that racial differences (even racial superiority) could have any bearing on questions of cosmopolitical right. As free beings with natural or human rights, human beings are all equal: the members of one race have exactly the same rights as the members of another. A more civilized race has no right to enslave, to dispossess, or to impose its civilization on a less civilized race. And Kant never wavers from his severe condemnation of Europeans for behaving as if it were otherwise.

index

a priori cognition, 24–32, 35–8
acceptance (*Annehmen*), 182–3
actuality (*Wirklichkeit*), 55, 95
Adams, R. M., v, 82–3
aesthetic ideas, 152, 165, 167–9
aesthetic judgment, 151–70
Africa, 178
agriculture, 116–17
Allison, H., 44, 82–3, 109, 169
Ameriks, K., 45, 109
Analogies of Experience, 57–61
analytic judgments, 25–7, 45
anthropology, 2, 9, 98–100, 110–27, **129**, 187–8; practical, 132–4
Anticipations of Perception, 57
antinomies, 89–100
appearances, 63–76
apperception, 33, 50
apprehension, synthesis of, 48
Aquila, R., 82
architectonic, 79, 108–9, 151
Aristotle, 39, 41
art, 164–9
assent, 182–3
association of ideas, 49
astronomy, 6, 29
autonomy, 131, 141–3; of the aesthetic, 160–2
Axioms of Intuition, 57

Baron, M., 150
Baumgarten, A. G., 16, 154
beauty, 151–62; and morality, 160–2; empirical interest in, 161; free and adherent, 158–9; rational interest in, 161–2
Beck, L. W., 23, 150
Beethoven, L., 8
Beiser, F., 23, 82
Bennett, J., 109
Berkeley, G., 10, 29, **60, 63, 67, 69, 71–2,** 73, 76
biology, 115, 126
body, 26, 67
Boileau, N., 162
Bok, H., 109
Brahms, J., 23
Burke, E., 162

Cambridge Platonists, 152
Canon of Pure Reason, *99, 108*
Cassirer, E., 23
casuistry, 131
catechism, 4
categorical imperative, 130–2, 135–43
categories, 41–4, 46–57
Catherine the Great, Empress of Russia, 176
causality, 26, 40, 41, 56, 57–8, 64, 78, 79, 80, 91, 96–100, 113, 142–3, 180
chemistry, 2, 5–6
chiliasm, 112
China, 178
Christianity, 18–20, 185–6
church, 20, 183–6
Cicero, 68
cinnabar, 49, 51, 54
citizen, active and passive, 175
civil society, 118–19
civilization, 118–19, 178
Clarke, S., 35, 101
classicism, 166–7

coercion, 144–5, 171–7
Cohen, T., 169
communication, 156, 166; freedom of, 134; public, 107–8
community (*Gemeinschaft*), 58, 60
composition (*Zusammensetzung*), 91–2
concepts (*Begriffe*), 30–2, 54–7; of reflection, 73
condition, 89–91
Confucius, 5
Constant, B., 187
constitution, civil, 118–21
continuity, principle of, 80–1
Copernican Revolution, 1, 29, 53, 70
Copernicus, N., 29
cosmological proof, 102
cosmology, rational, 76, 79, 89–100
Crusius, C. A., 28, 70, 89
Cummiskey, D., 150

Darwin, C., 125–6
democracy, 175
DePierris, G., 45
Descartes, R., 24, 28, 30, 50, 60, 61, 102–6, 152
despotism, 175
dignity (*Würde*), 131; *see also* autonomy; end in itself
discipline of pure reason, 107–8
division of labor, 117
dreaming, 39, 59
duties, 120; ethical, 140, 143, 145–50, 172–3; juridical, 143–5; of virtue, 146–50; perfect and imperfect, 138, 146; system of, 143–50; to oneself and to others, 146
duty (*Pflicht*), 129–30

Eliot, T. S., 48
Ellington, E. K. ("Duke"), 160
emotivism, 104–5
empiricism, 25, 82; about taste, 154, 157
end in itself, 139–41
Enlightenment, the (*Aufklärung*), viii, 2, 4–5, 13–14, 102, 108–9, 179–80
ens realissimum, *see* God
enthusiasm (*Schwärmerei*), 11, 73, 101, 152, 164, 186
equality, human, 140
Erlangen, 8
ethical community, 20, 183–6
ethics, 7–8, 14–15; *see also* duties, ethical; morality
Euclid, 38
European Union, 126, 177
evil, radical propensity to, 20, 118, 127, 133, 183–6
exceptions to moral rules, 130–2, 136–7
executive power, 175
experience (*Erfahrung*), 47–54, 58–61

faith: ecclesiastical, 183–6; moral, 108, 113, 120–4, 180–3
Falkenstein, L., 45
Feder, J. G., 10, 17
federation of states, xi, 120–1, 123–7, 177–9
Fichte, J. G., 12, 16, 32, 82, 112, 150, 169
form of purposiveness, 157
formalism: in aesthetics, 158; in ethics, 130, 139
Formula of Autonomy (FA), 135, 141–3
Formula of Humanity as End in Itself (FH), 135, 139–41, 146
Formula of the Law of Nature (FLN), 135, 136–9, 147
Formula of the Realm of Ends (FRE), 136, 141–3
Formula of Universal Law (FUL), 135, 136–9, 146
Förster, E., ix, 23, 61
France, 21, 177
free play, 155
free will, 25, 40, 78, 80, 91–2, 94, 96–100, 142–3
freedom: external, 20, 107–8, 173–7; practical, 96–100; transcendental, 96–100
French Revolution, 21, 176
Friedman, M., ix, 44
Friedrich II (the Great), King of Prussia, 4–5, 17, 176, 179

Friedrich Wilhelm I, King of Prussia, 5
Friedrich Wilhelm II, 17, 18, 21–2
Fundamental Principle of Morality, 134–43
Funk, J. D., xi, 8, 11

Gardner, S., 44, 83
Garve, C., 10
Gawlick, H., 23
Geach, P., 68–9, 83
genius, 166–9
geometry, 38
God, 6, 25, 40, 71, 76, **78, 79, 81, 89,** 100–6; his intuitive **understanding,** 31–2
Godwin, W., 11
Goethe, J. W., 23
Göschel, J., 11
Green, J., xi, 6, 8, 9, 10, 12, 13
Grier, M., 109
Guyer, P., viii, 23, 109, 150, 152, **169–70**

Halle, 8
Hamann, J. G., 11
happiness (*Glückseligkeit*), 141–3, 147–8, 180–3
Hare, J., 186
Hassidism, 4
Hegel, G. W. F., xii, 16, 32, 112, 119, **128,** 133, 169
Henrich, D., 62
Heraclitus, 68
Herbert, M., 11
Herder, J. G., 12, 13, 16, **159**
Herman, B., 150
highest good, 20, 180–3
Hill, T., 150
Hippel, T. G., 10–11, 13
history, 13–14, 20, 98–9, 108–9, 110–28, 130, 133, 139, 143; economic basis of, 116–17
Hobbes, T., 18, 30, 176
homogeneity, principle of, 80–1
Hume, D., xii, 4, 6–7, 16, 23, 25, 30, 48, 49, 62, 83, 101, 154, 158
Hutcheson, F., 4, 7, 16, 154, 160
hypothetical imperative, 136

Ideal of Pure Reason, 100–6
idealism, 60–1, 73; *see also* transcendental idealism
ideas, aesthetic, 165, 167–9
ideas of reason, 76–82, 84–107, 112, 113–14, 168
identity, 64–5, 68–70, 83
illusion, transcendental, 40, 76–8, 84–6
imagination, 55, 170
immortality, 25, 40, 81, 86–9
imperialism: American, 178; European, xi, 178–9
inclination, 129–30, 133
indifferentism, 25
Indonesia, 178
infinite judgment, 42
innate knowledge, 28–9
intelligible world, *see* things in themselves
intuition (*Anschauung*), 30–1, 35–**9, 54,** 85, 105–6; pure forms of, 35–9
Iraq, 178

Jacobi, F. H., 17, 73
Jacobi, J. K., 11
Jacobi, M. C., 11, 23
Jacobs, B., 127
James, W., 182
Japan, 178
Jena, 8
Jesus, 20
Joseph II, Emperor of Austria, 176
judgment (*Urteil*), 33, 42–4, 50–4; determining and reflecting, 155
Juno's peacock, 168
Jupiter's eagle, 168
justice, 118–19; *see also* right

Kain, P., 127
Kalinin, M., 3
Kames, Lord, 4
Kant, I.: character, ix–xii; daily schedule, 12–13; diet, 13; dogmatic slumber, 6–7; friendships, xi, 10–12; historical influence, 1–2; house, 12–13; life, 2–23; opposition to colonialism, xi, 178; racism, xi, 187–8; refusal to participate in religious services,

Kant, I. (*cont'd*)
19–20; works: *An Answer to the Question: What is Enlightenment?* (1784), 13–14, 108, 130; *Anthropology from a Pragmatic Standpoint* (1798), 22, 110, 127, 134, 186, 187–8; *Conflict of the Faculties* (1798), 22, 108, 110, 114, 130; *Conjectural Beginning of Human History* (1786), 13, 110, 116–17, 134; *Critique of Practical Reason* (1788), 15, 20, 96, 113–14, 118, 122, 129, **141**, 142, 149, 151, 181; *Critique of Pure Reason* (A:1781, B: 1787), 9–10, 14–15, 24–109, 134, 142, 151, 180; *Critique of the Power of Judgment* (1790), 16, 110, 112, 114, 119, 122, 132, 133, 151–70, 171, 181; *Groundwork for the Metaphysics of Morals* (1785), 14–15, 20, 96–7, 108, 129–43, 145, 146, 151, 171; *Idea for a Universal History With a Cosmopolitan Aim* (1784), 110–24, 132; *Lectures on Logic* (ed. Jäsche) (1800), 22; *Lectures on Pedagogy* (ed. Rink) (1803), 22; *Lectures on the Philosophical Doctrine of Religion* (c. 1783, ed. Pölitz, 1817), 101; *Lectures on Physical Geography* (ed. Rink) (1800), 22; *Metaphysical Foundations of Natural Science* (1786), ix; *Metaphysics of Morals* (1797–8), 15, 22, 129, 132, 134, 140, 141, 143–50, 171–9; *New Elucidation of Metaphysical Cognition* (1755), 6; *On A Presumed Right to Lie from Philanthropy* (1798), 130, 187; *On the Common Saying: That May Be Correct in Theory, But It Will Not Work in Practice* (1793), 111, 130, 176; *On the Distinctness of the Principles of Natural Theology and Morals* (1764), 7; *On the Forms and Principles of the Sensible and Intelligible World* (1770), 9; *Opus Postumum* (1798–1803, ed. Adickes, 1936, 1938), ix, 23; *Prolegomena to Any Future Metaphysics* (1783), 6–7, 10; *Religion Within The Boundaries of Mere Reason* (1794), 19–20, 96, 118, 130, 133, 179–86; *The End of All Things* (1794), 130; *The Only Possible Ground of Proof for a Demonstration of God's Existence* (1763), 101–2; *Toward Perpetual Peace: A Philosophical Project* (1795), 21, 110, 120, 124, 126, 130, 176–9; *Universal Natural History and Theory of the Heavens* (1755), 6; *What Does It Mean To Orient Oneself in Thinking?* (1786), 108, 130, 134
Kant, J. G., 3
Kanter, J. J., 10
Kaufman, A., 186
Keller, P., 62
Kemal, S., 170
Keyserlingk, Count and Countess, 5–6, 13
Kiesewetter, J. G., 18
Kitcher, P., 44, 62
Kitcher, Ph., 45
Kleingeld, P., 128
Knutzen, M., 5, 89
Königsberg, 3–4, 8–9, 20
Korsgaard, C., 109, 150
Kraus, C. J., 12
Kriemendahl, L., 23
Kuehn, M., 23
Kunz, D., 128

La Place, P.-S., 6
Lampe, M., 12
Langton, R., 82
latitude (*Spielraum*), 146–50
Lavoisier, A., 1, 22
law: positive, 173; Roman, 47
League of Nations, 126
legislative power, 175
Leibniz, G. W., 5, 25, 26, 28, 30, 35–7, **50**, 69, 73, 76, 89, 101, 103, 152
liar paradox, 99
Locke, J., 28, 30, 50, 117, 174
logic, 39–42, 78–9, 84–6
Longinus, 162
Longuenesse, B., 44, 45, 62
Louden, R., 150
Louis XVI, King of France, 176
lying, 130, 131, 187

magnitude, 57
Maimon, S., 17
Malebranche, N., 89
mark (*Merkmal*), 26
Marx, K., 112, 117, 126
masturbation, 130
materialism, 89
mathematics, 25, 26–7, 35–9, **40, 107**
matter, 40, 60–1, 67
maxims, 135–9
McFarland, J., 127
Melnick, A., 62
Mendelssohn, M., 7, **16, 100, 111,** 154
meta-critique, 32, 44
metaphysical deduction, 41–4
metaphysics, 24–7, 76–82
method, transcendental doctrine **of,** 106–9
Methodism, 4
mind–body problem, 88–9
mirage, 84–5
Montaigne, M., 117, 127–8
Montesquieu, Baron, 175
moral argument for God's existence, 108, 180–3
moral law, 134–43
morality, 20, 129–50, 151; and **beauty,** 160–2
Motherby, E., 11
Motherby, J. B., 8
Mozart, W., 160
Mulholland, L., 186
Müller-Lyer illusion, **84–5**
music, 159–60

nativism, 28–9
Neiman, S., 82
Newton, I., 35–7, 166
noumena, *see* things in themselves
numinous, 164

objectivity, 32–5, 48, 50–7
Ockham, W., 81
ontological proof, 102–6
original contract, 176
Orwell, G., 67
Otto, R., 164

paralogisms, 86–9
passion (*Leidenschaft*), 117–18
Paton, H. J., 150
peace, perpetual, 21, 119–26; *see also* war
perfectibility, 116
perfection, 147–8, 157
Perry, J., 83
perspectivity, 32–5
phenomena, *see* appearances
phenomenalism, 72
physical geography, 2, 6
physical influence, 89
physicotheological proof, 102
pietism, 4–5, 7
Platner, E., 9, 114
pleasure, disinterested, 155, 161–2
Plessing, F. V. L., 11, 23
political state, 118–19, 171–9
politics, 20–1, 171–9
possibility, ground of, 101
Postulates of Empirical Thinking, 57, 95
Posy, C., 83
pragmatism, 179
predisposition (*Anlage*), 115
pre-established harmony, 89
preformation, 28; *see also* nativism
priestcraft (*Pfaffentum*), 183–6
principles (*Grundsätze, Prinzipien*), 77–8, 174–5; *see also* reason
Prussia, 3–5, 19–22, 175, 177
psychology, rational, 76, 79, 86–9
publicity, principles of, 177
purposiveness without an end, 157
Putnam, H., 62

Quakerism, 4
quid iuris, 47
Quine, W., 45
Quinn, P., 186

races, 187–8
rationalism, 7, 30; about taste, 154, 157
Rawls, J., 186
realm of ends (*Reich der Zwecke*), 129, 136, 141–3, 183–6

reason (*Vernunft*), 77–82, 84–109, 116; polemical use of, 107; theoretical and practical, 122–4, 151–2
reciprocity (*Wechselwirkung*), 58
recognition, synthesis of, 50
refutation of idealism, 60–1, 73
regulative principles, 80–2, 96, 120
Reinhold, K. L., 15, 32
religion, xi, 19, 21–2, 101–2, 151, 152, 164, 179–86
reproduction, synthesis of, 49
republicanism, xi, 21, 175
Reuter, A. R., 3
right (*Recht*), 171–9, 187; cosmopolitan, 177–9; of nations, 177–9; of property, 174–5; principle of, 144–5; private, 174–5; public, 174–7
rights, moral, 172
Rosen, A., 187
Rossi, P., 187
Rousseau, J.-J., 7, 12, 21, **116, 117, 127,** 174, 175, 177

Saint-Pierre, Abbé, 21, 177
Salieri, A., 160
Schelling, F. W. J., 16, 22, 169
schema, 54–7, 88
Schiller, F., 3, 133
Schleiermacher, F., 185
Schulz, F. A., 4
science, 37–8, 99, 159
scripture, 185
self-affection, 82
self-conceit (*Eigendünkel*), **118, 133,** 150
Selle, C. G., 17
sensibility (*Sinnlichkeit*), 31, 35–9
sensible world, *see* appearances
Sextus Empiricus, 33
skepticism, 25, 33–6, 44, 46–7, 56–7, 60–1, 89
Smith, A., 4, 12
Socrates, 25
sophisma figurae dictionis, 87
soul, 25, 76, 81, 86–9; *see also* immortality
space, 34, 35–9, 61, 91
specification, principle of, 80–1
Spinoza, B., 30, 73
spirit (*Geist*), 167
Stalin, J., 3
Sterne, L., 10
stoicism, 133
Sturm und Drang, 11
sublime (*Erhabene*), 162–4
substance, 40, 41, 57, 78, 86–**9, 91**
suicide, 130–1
supernatural assistance, 89
superstition (*Aberglaube*), 100, **179–80,** 183–6
syllogisms, 78–9, 85, 86–7, 175
symbol, 161–2
synthesis, 47–54
synthetic judgments, 26–7, 45

taste, 151–70; judgments of, 153–60
taxonomy, 80–1
teleology: moral, 146–50; natural, 113–16, 120–1, 124–7, 151–2, 169
Teske, J. G., 5
theology: minimum of, 182; rational, 76, 79; *see also* God
things in themselves, 29, 63–76, 93–6, 142–3; *see also* transcendental idealism
thinking, 30–1, 54–7
time, 34, 35–9, 47–50, 58–61, 80, 90, 91–2, 97–100
time-determination, 58–61
transcendental aesthetic, 35–9, 106
Transcendental Analytic, 40
transcendental deduction, 34, 41, 46–**54,** 55, 58
transcendental dialectic, 40, 76–82, 84–109
transcendental idealism, 27, 29, 38–9, 61, 63–76, 93–6
transcendental logic, 39–44, 106
transcendental philosophy, 24, 39
true and immutable natures, 105–6
truth, 53–4, 62
Tully, *see* Cicero
Turgot, A., 119, 128
two standpoints, 99–100
two worlds, 65–6

Ulpian, 173
unconditioned, 89–100
unconscious representations, 50
understanding (*Verstand*), 31, 39–**44**, 54–7, 155
United Nations, 126
unsociable sociability, 117–18

Velkley, R., 127
virtue (*Tugend*), 149–50
Voltaire, 112

war, 20–1, 119–20, 178–9
Waxman, W., 45
Welk, L., 160
Wieland, C., 15
Williams, H., 187
Wizenmann, T., 17
Wolff, C., 5, 6–7, 78, 101
Wöllner, J. C., 17–19, 22, 179
Wood, A., ix, 109, 150, 187
world, 76, 80, 89–100
Wreen, M., 187

Yovel, Y., 127

Zedlitz, Baron von, 17
Zeno, 99

"Truly up-to-date overviews of Kant's critical philosophy are rare. Allen Wood's outstanding new study is at once an intriguing philosophical interpretation of Kant for specialists and an extremely clear and helpful guide for students."

Karl Ameriks, University of Notre Dame

"Allen Wood's *Kant* is clearly written, and it is accessible to a broad audience without making any sacrifice of precision in expounding complex philosophical doctrines. It covers all aspects of Kant's thought and relates them, as much as is possible within the constraints of a fairly short book, to contemporary issues and concerns."

Béatrice Longuenesse, New York University

"Writing with both vigor and rigor, Allen Wood provides concise and insightful accounts of Kant's theory of knowledge and critique of traditional metaphysics, and of his moral and political philosophy. He also paints a distinctive picture of the historical dimension of Kant's thought: his sensitivity to the history of philosophy, but even more importantly his recognition that humankind itself has a history, and that its moral goals must be achieved through the means of history and within the limits of history."

Paul Guyer, University of Pennsylvania

This lucid survey takes readers on a thought-provoking tour through the life and work of Immanuel Kant. The book provides expositions of Kant's key works, especially his *Critique of Pure Reason*, and represents the broad range of his philosophical thought, including his contributions to metaphysics, epistemology, the philosophy of science, the philosophy of history, moral philosophy, aesthetics, politics, and the philosophy of religion.

The first volume in the *Blackwell Great Minds* series, Allen W. Wood's *Kant* offers students, scholars, and interested readers new insights into one of history's most important and influential philosophers.

Allen W. Wood is Ward W. and Priscilla B. Woods Professor at Stanford University. He is the author of seven books, including *Hegel's Ethical Thought* (1990), *Kant's Ethical Thought* (1999), *Unsettling Obligations* (2002), and *Karl Marx* (second edition, 2004). He is co-editor, with Paul Guyer, of the *Critique of Pure Reason* (1999).

Cover illustration: anonymous portrait of Kant, c.1790. Photo AKG Images.

Cover design by cyandesign.co.uk

Printed in the United Kingdom

Visit our website at
http://www.blackwellpublishing.com

Copyrighted material